CITeR TECHNOLOGY ROADMAP 2022

CITeR
Center for Identification Technology Research

FORWARD

As CITeR passes its 20th year, it is clear that challenges remain in identification technology research. There is rapid deployment of biometric and identity technologies from far reaching applications including border security, defense, benefits distribution, forensics, and disaster relief, to consumer electronics, e-commerce, and banking. This expansive use makes it clear that continued research is needed to address the unique range of technological challenges including sensing, quality, matching, fusion, security, privacy, perception, fairness, and distinctiveness.

To aid in methodically addressing the needs, CITeR launched the Technology Roadmap effort. The goal of the CITeR Technology roadmap is to establish a resource for utilization by researchers and government and industrial organizations to help guide and develop research toward achievement of a future state.

The creation process is fueled by an exchange between researchers and government and industrial affiliates working together to develop end state vision and operation concepts along a 3-10 year timeline. The resulting technology roadmap is envisioned as a living reference document to help guide CITeR's work for years to come.

CONTRIBUTORS

Jeremy Dawson, Brian Greene, Daqing Hou, Xin Li, Xiaoming Liu, Siwei Lyu, Sebastien Marcel,Dan McCaugherty, Nasser Nasrabadi, Terry Riopka, Arun Ross, Stephanie Schuckers, Ranga Setlur, Matt Valenti, Wenyao Xu

CONTENT

1. Approach

The effort began with a small core team of government, academia & industry members followed by a broadened effort to receive feedback from the larger CITeR community. The framework has two tiers: (Tier 1) Application Areas and (Tier 2) the Technology Concepts. Application Areas give the future vision and Technology Concepts describe the enabling technologies needed to achieve that vision.

a. Application Areas (Tier 1)

The Application Areas are how we envision identification technologies in the future. The approach is blue-sky thinking which is not constrained by today's technology. Thisincludes operational concepts and advanced system architectures. The output for each is a 1-2 page description detailing the applications including 'Current Mission Needs'and 'Enabling Technologies'.

A. Biometrics in Border Control
B. Standoff Identification
C. Trust in biometrics
D. Real-time Threat Identification
E. Identity Proofing
F. Forensic Analysis
G. Encounter-Based Intelligence Analysis
H. Consumer Authentication and Access Control

b. Technological Concepts (Tier 2)

The next step is determining the enabling Technology Concepts needed to realize Application Areas. The Technology Concepts are mapped to Application and grouped into seven Technology Concepts "Families" listed below. Each Technology Concept has a 1-2 page description containing Background, Future Vision and Keywords.

1.0 Sensors and data collection
2.0 Image quality, preprocessing
3.0 Biometric modality matching
4.0 Data analytics & fusion
5.0 Security, privacy, perception
6.0 Distinctiveness, permanence, demographics
7.0 Other, extension to another area beyond biometrics

c. Identify strategic technology progression

The last step is to identify most impactful technologies that require research and development and establish a time-phased investment priorities based on near-term, mid-term, and far-term, as defined below.

Near Term:

- Proven in laboratory or test environments using operational conditions – it is mature.
- Rapidly emerging technology with broad investment – it will develop quickly.
- Does not require many organizations to determine acceptability of solutions or pre-coordination has occurred with general agreement of acceptability of solutions – ready to gain acceptance.

Mid Term:

- Proven in a research environment, more technology development needed. Not evaluated under operational conditions.
- Niche technology without broad investment.
- Requires coordination across many (3+) organizations to determine acceptability of outcomes.

Long Term:

- Technology concept is notional or theoretical. May depend on technology that does not yet exist.
- Very little research has been performed. Research in disparate areas may be needed to make technology progress.
- Technology development will require many years of coordination across multiple organizations (3+).

Biometrics in Border Control

	NEAR-TERM 0-5 YEARS	MID-TERM 6-15 YEARS	FAR-TERM 15+ YEARS
SENSORS & DATA COLLECTION | 1.0	- Higher quality contactless fingerprint capture	- Group face capture at speed [1.1]	- Automobile-based biometrics & people counting; high quality face capture through moving automobiles, automobile identification
IMAGE QUALITY, PREPROCESSING | 2.0		- Biometrics on the move, outdoors [2.2]	
BIOMETRIC MODALITY MATCHING | 3.0	- Resource constrained algorithms [3.1]	- Better facial recognition algorithms which matching with partially masked faces [3.1]	
DATA ANALYTICS & FUSION | 4.0			
SECURITY, PRIVACY, PERCEPTION | 5.0	- Resistance to high tech attacks, deep fakes, morphing, doppelgangers [5.1]	- Secure biometric data sharing [5.2]	- Advanced privacy protections for the storing and sharing of biometric data [5.2]
DISTINCTIVENESS, PERMANENCE, DEMOGRAPHICS | 6.0	- Differentiation of twins and doppelganger - face recognition for high resolution imaging [6.3]		
OTHER, EXTENSION BEYOND BIOMETRICS | 7.0	- Digital document design, issuance, verification, authentication		- Threat assessment: contraband, weapons, backpacks, bombs, etc - MRI/mmWave Scan

2.2 Face Pose Invariant Features and Face Frontalization

Unconstrained face recognition at extreme pose is an important and a challenging problem. The existing methods that address this pose problem can be generally categorized into two major categories. The first category aims to obtain pose-invariant embeddings [1, 2, 3]. The other aims to normalize (frontalization) face images [4, 5, 6, 7] to identity-preserved frontal views (as a preprocessing step), which can be directly used by any off-the-shelf face recognition systems without the knowledge of the recognition models. For the first category, deep metric learning [1] is a common way to achieve pose-invariant embeddings in a latent subspace. Due to the imbalanced distributions that characterize a long tail distribution of large pose faces, it is often difficult to achieve ideal pose-invariant embeddings across large pose variations. The second category is known as face...

For more information click here

citer.clarkson.edu

A. Biometrics in Border Control

Establishing the identity of the traveling public at the country's border is paramount. Biometric solutions are increasingly used to identify the traveler at multiple points throughout the travel and inspection process to provide a more personalized travel experience and increase the efficiency and security of the border control system. Many countries and political unions have different environments for border control such as airports, land ports, and seaports which support different modes of transport (i.e. auto crossings, pedestrian queues, bus processing). The mission of border control requires rapid, non-obtrusive biometric capture. The quality of the biometric samples must also be of sufficient quality to be searched and matched against large enterprise level biometric identification systems.

Many of the varied environmental scenarios for processing of travelers at the border present opportunities for the use of different biometric solutions. For example, some environments are outdoors, some require the use of mobile technology, and some are cooperative. These different scenarios invite out-of-the-box thinking of how to approach the problem and are open to new modalities and hardware for both enrollment and capture.

Enabling Technologies

- Privacy protections for the storing and sharing of biometric data
- Secure biometric data sharing
- Better facial recognition algorithms which matching with partially masked faces
- Differentiation of twins and doppelganger - face recognition for high resolution imaging
- Higher quality contactless fingerprint capture
- Mobile ID verification
- Remote enrollment of verified high quality biometrics
- Automobile-based biometrics and people counting; high quality face capture through moving automobiles, automobile identification (?)
- Biometrics on the move, outdoors
- Group face capture at speed
- Threat assessment: contraband, weapons, backpacks, bombs, etc

(Enabling Technologies continued)

- Resistance to high tech attacks, deep fakes, morphing, doppelgangers
- Digital document design, issuance, verification, authentication
- External API to clients for services: proliferation of services and APIs to simplify and standardize transactions within a biometrics domain architecture according to message standards
- MRI/mmWave Scan

B. Standoff Identification (long distance, 100 meters or more)

Face recognition at a distance (FRAD) is concerned with the automatic recognition of non-cooperative subjects over a wide area. This remote biometric collection and identification problem can be addressed with an active vision system where people are detected and tracked with wide-field-of-view cameras. Face recognition at a distance will enable watch-list recognition for security at terminals and critical infrastructure, intruder detection and identifying whitelist personnel. Furthermore, the ability to collect biometrics at a distance (e.g., aerial, satellite imagery) would provide tactical users with critical information on subjects without the need for their cooperation. Standoff biometrics capabilities may reduce the time on-site for tactical forces, minimizing the window of opportunity for hostile forces to ambush, maneuver, or collect intelligence on friendly forces. Major challenges in FRAD are: 1) detecting and obtaining sufficient face image resolution at long range during at anytime of day, 2) removal of image distortion caused by atmospheric effects, 3) tracking subjects and accurately matching facial images obtained from unconstrained, non-ideal poses. Therefore, the components of a high-performing FRAD system will consist of many different algorithms (image enhancement, super-resolution, key-frame selection, face detection, face tracking and recognition). Novel sensors such as multispectral/ hyperspectral and stereoscopic binocular imagery have been implemented to address these challenges. For example, hyperspectral cameras are known to provide better discriminative information for object recognition than visible spectrum ones. Hyperspectral imagery can be used for material classification and identification. However, a major problem with hyperspectral cameras is the insufficient spatial resolution for distance object/face recognition. Stereoscopic binocular imagers have been used to isolate persons of interest in images with several faces captured at a distance.

Examples of current mission needs:

1. **Face Detection:** Face detection at a distance in a scene is a major task since the quality of surveillance images are poor (i.e., insufficient pixels on the face). Face detection algorithms need to be investigated and combined with other image processing tools (e.g., image enhancement and super-resolution).
2. **Face Tracking:** A FRAD systems needs to be equipped with an adaptive target selection mechanism (a face in a watch list) based on the current actions and history of each tracked subject to help ensure that facial images are captured for all subjects in view for post-processing.

3. **Super-resolution:** Face recognition at a distance with low-resolution face images is a difficult problem in face biometrics. Advanced video super-resolution algorithms need to be developed and optimized solely for face recognition at a distance.
4. **Soft biometric information:** Soft biometric information extracted from a human body (e.g., height, gender, skin color, hair color, gait, and so on) is ancillary information easily distinguished at a distance, but it is not fully distinctive by itself in recognition tasks. However, this soft information can be explicitly fused with biometric recognition systems to improve the overall recognition when face images are captured in poor quality conditions with high variability.
5. **Multispectral/Hyperspectral face detection/identification:** Hyperspectral cameras are known to provide better discriminative information for object recognition than visible spectrum ones. For example, spectral measurements in infrared allow sensing of objects at night, and extract features which are invariant to orientation and reflection. Hyperspectral imagery can also be used for material classification and identification. Major challenges with hyperspectral cameras is insufficient spatial resolution for distance object/face recognition and cross-spectral face recognition (i.e. matching a short-wave infrared probe against a visible gallery).

- Sub-pixel personnel detection at an extreme distance: Aerial pixel/subpixel face/personnel detection is an important research area in tactical applications. Images captured from UAVs/satellites can be used for personnel face detection. Thermal hyperspectral imagery can be used to detect subpixel faces, because a single hyperspectral pixel contains multiple materials. Novel unmixing methods can be used to decompose a hyperspectral pixel into its endmember spectra.

C. Trust in Biometrics

Biometrics technologies have become the primary tools for verifying identity in recent times across many different applications from customs and border protection, to controlled access to facilities, to user authentication in smart devices. However, its continued acceptance requires addressing the issue of people's trust in the system which may be dented by anecdotal reports of issues such as compromised templates due to data breaches, or biases in the system, or performance shortcomings. Efforts to maintain trust will necessarily need to address three main objectives for the technology, viz, privacy of the biometric data, fairness of the system, and adoption of better standards to evaluate the "trustworthiness" of the technology.

Examples of current mission needs:

1. **Explainable biometric systems** that will enable introspection of models to ascertain the reasoning involved in the predictions from biometric features and increase transparency.
2. **Creating bias free representations**
3. **Data biases,** arising from the data distributions that biometric models are trained on, will need to be identified and rectified by explicitly accounting for the imbalanced data and implementing novel optimization algorithms that minimize its effect.

(Examples of current mission needs continued)

4. **Model/algorithmic biases** will need to be addressed and suitable changes in architecture and design implemented to improve the general fairness of the system.
5. **Increased transparency** on how the system manages user data and storage, while preserving the privacy of biometric templates to provide better assurance to users.
6. **Advances in cryptography** to facilitate the generation of highly secure biometric data that will be less prone to leakage of uniquely identifiable information in addition to being resilient to various types of presentation attacks..
7. **Providing increased cancelability** to biometric templates to promote non-linkability of the user's biometric data across various databases which addresses both issues of spoofing and user data leakage.
8. **Identification of current shortcomings** of biometric technology with respect to credibility will prompt creation of improved standards and metrics to evaluate performance and consequently, their reliability.
9. **Evaluation benchmarks** across various dimensions of fairness will result in clearer perceptions of how different biometric modalities and models perform.
10. **Revised standards** can describe expectations from biometric systems and levels of acceptability in various application areas which will reinforce positive public perception of such systems.
11. **Improvement in real time multimodal biometric systems** could reduce reliance on one particular modality of biometric recognition (like face). This would considerably reduce false positives in scenarios such as surveillance, thereby reducing misidentifications thereby improving trust.

D. Real-time Threat Identification

Within industrial settings, government facilities, or other controlled access locations, biometrics can play a role in rapidly identifying persons that can represent a potential threat through the identification of persons on a watchlist, persons of interest, or persons that do not have authorization for access to a controlled area. Future surveillance systems will have the capability to alert security personnel in near real time (within a few seconds) when a person of interest (POI) is detected using biometric technology. Such a surveillance system should also be able to detect obfuscation attacks, where an individual deliberately seeks to avoid being detected by it.

The identification of a threat would use technology such as facial recognition, contactless fingerprint collection, and other physiological characteristics such as gate, thermal characteristics, facial expression, threatening behavior (e.g., wielding a firearm) or a fusion of characteristics.

Collection technology will enable the collection of usable biometrics from longer distances as a result of higher resolution cameras and intelligent construction of usable data (e.g., creating a frontal face image from non-frontal images).

These surveillance eco-systems will comprise mobile and fixed cameras, access control biometric enrollment systems, watchlist systems, biometric matching systems, security operations centers, and alerting systems. The surveillance eco-system will increasingly move processing to the edge as

computing technology evolves providing faster local response to threat detection. Networks of surveillance systems within an enterprise will be able to share biometric data for potential threats, track POIs as they appear in different locations, and provide the capability to search live surveillance data such as video feeds to identify past encounters of a POI. Central machine learning systems will collect images and video of threatening behavior or situations to develop deployable models for automated identification of threats such as performing physical harm, exposing weapons, suspicious behavior, etc.

With the integration of GIS, security officers will track a POI's location history to develop and study patterns of behavior. On a broader scale, to include integration within national security systems, it will become easier to identify and track POIs.

Enabling Technologies

- Raising TRL levels for biometric detection from video feeds
- Development of smart edge devices for collection and alerting
- Edifice or campus level surveillance eco-systems will manage biometric data, matching, and detection of biometrics (e.g., faces)
- Enterprise level integration of data and systems

E. Identity Proofing

Identity proofing or establishment of identity for a relying party is vital for many real-world transactions and interactions. In a digitally interconnected world where all manner of personally identifiable information about most individuals is easily available from a simple online search, identity theft (or people posing as someone else) during the process of account creation (e.g. banking, credit, passports, visa, etc) will continue to be an issue of serious concern.

The first step to setting up an identity may involve some of the following steps: (1) having users present documents (e.g. passports, drivers license) in person or more commonly remotely, (2) checking that the documents are valid and not forged, (3) checking that the information on the document corresponds to the authoritative source (e.g. DMV, Dept of State), (4) determining if the individual presenting the document matches the biometric information present on the document (e.g. selfie versus ID photo match); (54) searching amongst a biometric database to ensure that the person matches against their biometric (if available) and that the person does not have multiple identities (deduplication); and optional (6) read and verify the ID information if the document is embedded with an IC chip.

In the future, we envision a world where individuals have control over their identity information and can choose identity providers to vouch for their identity to other relying parties. These trusted parties can incorporate biometric technology in a secure manner where their identity is assured and cannot be "taken over" by those who may want to steal it. These identity providers are able to reveal as much information as needed, such as confirming age without revealing birthday, confirming address, providing full identity information if approved by individual (e.g., applying for loan), etc.

Enabling Technologies

- Biometrics at birth, for a lifetime
- Remote enrollment
- Mobile IDs
- Ensuring high quality samples, dealing with low quality samples
- Studies in children/infants
- Detection of document fraud (e.g., paper fingerprint)
- Trusted path to biometric sensor/camera
- Identity as a service
- Biometrics as a service
- Self sovereign identity
- Resistance to attack
- Large scale search for deduplication
- Studies of individuality/permanence
- Architectural security
- Mobile liveness
- Trusted enrollment stations
- Embedded IC (e.g., PUF - physically unclonable functions)

F. Forensic Analysis

Biometrics processes employed in forensic analysis applications often require man-in-the-loop functionality to allow an expert examiner to review and confirm or overrule the decision made by an automated recognition system. This approach can also be employed during booking (i.e. enrollment) to provide the sensor operator with feedback on the quality of the captured images prior to entry into an EBTS or other data exchange format. Most forensic collections of biometrics use certified stand-alone sensor hardware, and matching is performed by commercial algorithms. Forensic examiners have expressed interest in applying biometrics pattern recognition and image processing methodologies to the evaluation of crime scene samples, such as latent fingerprints, video camera footage, and bullet striations. Recently, cellphone and body-cam technologies are being used to capture opportunistic iris and fingerprint images as well as face images. The implementation of rapid DNA at booking also opens new avenues for the use of DNA as an investigative tool.

Examples of current mission needs:

1. **Lights-out Latent Matching –** Currently, matching of latent fingerprint impressions collected from crime scenes are manually matched against exemplar images by certified latent print examiners. While latent print workstations have been developed to guide examiners in this process, automated matching to narrow down the pool of potential matches has not been implemented. Further work needs done in this area to improve the effectiveness of automated matching of latent to livescan or legacy 10-print fingerprint images.
2. **Latent fingerprint Quality Index –** NIST has already established a quality index measure for contact-based and contactless fingerprints for use during the image capturing process. However, there is no well-known algorithm to estimate the quality for the latent fingerprints. There is a need to develop new algorithms to estimate the quality of latent fingerprints in order to reduce the load needed to review a large number of latent fingerprints by an examiner.

(*Examples of current mission needs continued)*

3. **Non-contact Fingerprint Interoperability –** The increased capability of smartphone technology has led to improvements in the quality of image capture. Because field agents are all issued cellphones, they would like to be able to use them to capture fingerprints, rather than carry a separate stand-alone fingerprint device. However, due to photometric distortion, lack of elastic deformation, and overall input image quality (motion blur, defocus, etc.) the interoperability of contact and fingerphoto fingerprints needs improvement.
4. **Database Image Management/Quality Control –** There are over 15,000 state and local law enforcement agencies in the US. While the FBI provides guidelines for these agencies to enter data in the NGI system, the specific procedural details for image capture at booking is left up to the agency itself. Because of varying funding across the US, this leads to a large degree of variation in on-site training and equipment used, which can lead to quality issues associated with image capture. Automated methods of image analysis prior to and after images have been entered into law enforcement databases are needed to ensure that images of adequate quality and type are utilized in searches.
5. **Flexible Addition of New Modalities –** NGI was designed with modularity in mind, allowing for the addition of new modalities into the booking record as needed. Rolling these changes into the platform, however, is nontrivial. New approaches may be needed for modalities such as voice, etc.
6. **Opportunistic Iris Recognition –** The ubiquitous nature of high-resolution camera technology offers the capability of extracting a useable visible iris pattern from a facial or periocular image. However, most iris recognition algorithms are designed to match near infra-red (NIR) images at relatively low resolutions. New methods of cross-spectral & cross-resolution iris recognition are needed.
7. **Post-Mortem Identification –** Physical degradation of anatomical structures can make it difficult to collect fingerprint or iris biometrics from deceased individuals. In addition, facial recognition often fails if the eyes are closed or pupils cannot be determined in the image. New methods of biometric capture and matching are needed to overcome these challenges.
8. **Opportunistic Face Recognition –** Security camera and other forensic video evidence may contain non-ideal face images (non-fontal pose, non-uniform lighting, etc.) captured at a distance. Many of the challenges of face recognition at a distance/in the wild also apply here.

G. Encounter-Based Intelligence Analysis

In some applications, eliciting the identity of an individual may become necessary based on the type of encounter. Here, the focus may be on a single individual or on multiple individuals. This would entail characterizing the nature of an encounter and then determining the identity of the individuals involved. When a matching identity is not present in the database, then a new identity profile has to be created dynamically and populated with the biometric, biographic and other descriptive attributes of the individual. Further, the database has to be periodically de-duplicated in order to merge multiple profiles pertaining to a single individual.

As biometric and surveillance sensors evolve, it will become necessary to update identity profiles to incorporate biometric information of an individual emerging from these novel sensors. Further, the

encounters of interest may vary with time. Effectively modeling the different types of encounters would be necessary for not only characterizing the nature of an encounter but to also detect it in real-time. The database used in these operations should allow for the grouping of identities and for developing connections between identities involved in an encounter.

When generating new identity profiles based on an encounter, it will be necessary to assess the quality of the biometric data being used to populate it. Poor quality data may lead to false matches or false non-matches in the future and, therefore, an automated procedure is needed to curate and ingest data into the system prior to using them. In addition, the surveillance system must embody the privacy laws of the local region thereby ensuring that biometric data is stored, accessed and transmitted in accordance to the prevailing law.

Artificial Intelligence is used to automatically identify patterns of adverse behavior such as recurring encounters associated with known offenders or known illegal activity.

H. Consumer Authentication and Access Control

The use of biometrics in consumer authentication and access control is progressing at a rapid pace. It has been driven by the desire to eschew passwords and tokens (something you know or have) in favor of methods that rely on your unique identity (something you are) to enable access to personal devices, systems and physical entry points. Consumer acceptance has accelerated due, in part, to the momentum gained through the recent pandemic, making non-obtrusive authentication more appealing as a result of its increased safety and convenience.

Research and development continues toward the goal of secure, accurate and frictionless authentication. Despite many advancements, a fast changing world is constantly pushing the limits of biometric technologies.

Examples of current mission needs:

1. **Advancements for existing biometric modalities**
   - improved face and voice authentication in unconstrained environments
   - improved recognition using video streaming and super resolution techniques
   - touchless fingerprint and iris-at-a-distance authentication
2. **Investigation of new modalities**
   - Multi-spectral (NIR, SWIR, LWIR), millimetre (MMW) and submillimetre (SMW) waves for generating new biometric signatures
   - Body contours
   - Ear prints
   - Finger/palm/wrist vein mapping
   - Gait analysis
   - Remote magnetoencephalographic scanners for detecting brain signatures

3. **Novel multi-modal biometric strategies**
   - Novel combinations of multiple modalities
   - Feature level fusion techniques
4. **Strategies for silent continuous authentication**
   - Continuous passive behavioral biometrics authentication
   - Context-based analysis of behavioral data
   - Murphy's Law for behavioral data: "whatever data can be collected, will be collected"
5. **Wearable, embedded or ingestible devices for bio-biometrics**
   - 'Natural body signature identification'
   - Subcutaneous implants
   - Neural implants
6. **Liveness/Presentation attack detections (aka Spoofing detection)**
   - New sensors/hardware/software for liveness detection to enable secure remote biometric authentication
7. **Cloud vs. Edge Computing Strategies for Biometric Authentication**
   - New ramifications of the Internet of Things on biometric authentication

Topics suggested to possibly include under this topic: External API to clients for services: proliferation of services and APIs to simplify and standardize transactions within a biometrics domain architecture according to message standards, Internal APIs to underlying subsystems: service based transactions and messaging standards.

Biometric marketplace: interoperable components within a biometrics domain architecture stimulating advancement of biometrics COTS products, testing harness to compare, selection of best-of-breed components for a particular problem, Privacy incorporated into the architecture, Architecture design: functional split between local device and cloud.

Concept Families

1.0 Concept Family: Sensors and Data Collection

1.1 Technology Concept: Facial Imagery

1.1.1 Background

Camera technologies have forever been the foundation of facial biometric systems. Initially used primarily for mugshots, best practices, such as the ANSI NIST SAP50/51 face photo guidelines, ensure that face images used in biometric facial recognition systems are captured with uniform lighting, neutral expression, and uniform backgrounds at specified resolutions. As facial recognition applications move increasingly into unconstrained applications, camera systems need to be able to capture face images in nonuniform (or no) lighting, at long distances, at variable resolutions, and a variety of other operational challenges. In tandem with new camera technologies and uses, research datasets comprised of facial imagery captured in unconstrained conditions are needed to conduct hardware performance evaluations and develop new matching algorithms. In the case of the ever-growing application of deep learning and artificial intelligence (AI), very large face datasets with images captured under a myriad of conditions are essential for training and testing phases to be successful.

1.1.2 Future Vision

Sensor Hardware: The future of facial biometrics will primarily involve facial capture in unconstrained conditions. Border control applications are seeking to perform face capture of large groups of individuals moving through checkpoints, either on foot or in vehicles. Smart camera systems that can track and capture faces, perhaps through the windows of moving automobiles, or perform automated person counting and frame segmentation, will be needed in these applications. Beyond facial capture, smart cameras with some level on on-board intelligence can potentially also provide encounter-based intelligence, such as determining the make and model of vehicles containing people, or the intent of individuals based on their behavior, gait, or mannerisms. Depending on the resolution and environmental conditions, these camera systems may be adapted to capture iris images at a distance as well.

One technology that could be adapted to meet this application need are foveal-vision cameras. Hierarchical-foveal-machine-vision (HFMV) systems [[1]] mimic the human retina by providing higher resolution in a region of interest, with decreased resolution in other areas. In these systems, resolution is essentially a dynamically allocable resource, allowing higher frame rates without increasing processing times. Such systems require an automated object tracking algorithm to 1) control a pan-tilt-zoom (PTZ) gimbal to ensure that the object of interest remains in the region of highest resolution (e.g. the center of the field of view/ sensor array) or 2) modify the CMOS array to adapt the resolution of a specific region within the camera field of view as the object of interest changes position. This technology was developed in the late 90s, but AI and deep learning technologies make HFMV cameras prime

candidates for application in current and future biometrics deployments to enable higher levels of encounter-based intelligence in standoff applications.

Encounter-based intelligence is not always available in daytime or other well-lit applications. Face capture at night using beyond-visible wavelengths ranging from near-infrared (NIR) to short-wave infrared (SWIR) to long-wave infrared (LWIR) has been a standard of night operations for many agencies, and biometrics approaches to cross-spectral face recognition have resulted in acceptable matching performance [[2]]. Despite recent success, further improvement in night-vision camera hardware is necessary to boost the resolution of LWIR imagers. Sub-pixel and sparsity-based detection of faces and other targets has been performed using visible hyperspectral imagery [[3]]. These techniques could be applied to LWIR-specific hyperspectral imagers [[4]] to improve the performance of LWIR sensing devices.

Far below the hardware complexity of these specialized imaging systems, CMOS cellphone cameras are being pressed into use by many agencies as a multi-biometric capture tool. Similarly, other CMOS platforms, such as mirrorless and SLR digital cameras, action cams, body-worn cameras, UAV/UAS-mounted systems, security cameras, and many others, are being used for facial biometric capture. However, just because the hardware can capture a face image does not mean that the resulting image is suitable for biometric recognition. To address the proliferation of camera hardware in modern society, government agencies are developing a set of hardware requirements for cameras to ensure only hardware approved for biometrics applications are used for facial recognition.

Unconstrained Face Datasets: To address the dearth of facial imagery captured in unconstrained conditions, agencies such as the Intelligence Advanced Research Projects Activity (IARPA) are including face image collection in their project tasking for efforts funded to support the development of AI approaches to facial recognition. Face image datasets collected for the now-completed Janus program, IJB-A, IJB-B, and IJB-C [[5]], contain tens of thousands of visible, SWIR, and LWIR face images and videos in unconstrained operational scenarios that have been used in NIST face recognition challenges and face recognition vendor tests (FRVT). One major component of the recently announced IARPA Biometric Recognition and Identification at Altitude and Range (BRIAR) program is the collection of face images and video at ranges up to 1000km and from elevated platforms, including security cameras and aerial sensor platforms [[6]]. Funded efforts will be expected to collect face and whole-body imagery from at least 800 individuals to support the development of advanced whole-body and face recognition algorithms. While these efforts are producing a significant number of samples that are suitable to train AI systems to handle unconstrained facial imagery, data collection will be a need for the foreseeable future. Such efforts can be costly, as the amount manpower needed to manually annotate and categorize the ground-truth image information can be immense. This manpower need can be met by using services such as Amazon Mechanical Turk, but trained biometrics practitioners may be better at annotation than the general public.

Keywords

• Smart cameras and sensors • Border control • Biometrics on the move
• Encounter-based intelligence • Real-time threat identification and detection
• Standoff identification • Face-specific cameras • Unconstrained environments

RETURN TO CHART

References

[1]. D.C. McKee, C. Bandera, S. Ghosal, and P.J. Rauss "Model-based automatic target recognition using hierarchical foveal machine vision", Proc. SPIE 2755, Signal Processing, Sensor Fusion, and Target Recognition V, (14 June 1996); https://doi.org/10.1117/12.243150

[2]. T. Bourlai (Ed), Face Recognition Across the Imaging Spectrum, Springer International Publishing, 2016.

[3]. W. Lv and X. Wang, "Overview of Hyperspectral Image Classification", Journal of Sensors, vol. 2020, Article ID 4817234, 13 pages, 2020.

[4]. D. Manolakis et al., "Longwave Infrared Hyperspectral Imaging: Principles, Progress, and Challenges," in IEEE Geoscience and Remote Sensing Magazine, vol. 7, no. 2, pp. 72-100, June 2019.

[5]. https://www.nist.gov/programs-projects/face-challenges

[6]. https://www.iarpa.gov/index.php/research-programs/briar/baa

---

2.0 Concept Family: Image quality, preprocessing

2.1 Technology Concept: Cross-spectral Face Matching and Synthesis

2.1.1 Background

In recent years, there has been significant interest in Heterogeneous Face Recognition (HFR) [1], where the goal is to match visible facial imagery to facial imagery captured in another domain, such as the infrared spectrum [2, 3], polarimetric thermal [4], or millimeter wave [5]. Since there is a significantly less facial imagery available in these alternative domains compared to the visible domain, a robust cross-domain facial matching cannot easily be achieved. In recent years, there has been growing research on thermal-to-visible face recognition [6,7,8] for night-time surveillance and low-light scenarios. Visible images contain rich textural and geometric details across key facial structures (i.e., mouth, eyes, and nose). However, in conventional thermal facial imagery, though some edges around the eyes and eyebrows do appear, but they suffer from significant lack of contrast compared to the

corresponding visible images, thus highlighting the large domain gap.

Algorithms for thermal-to-visible face recognition can be categorized as cross-spectrum feature-based methods, or cross-spectrum image synthesis methods. In cross-spectrum feature-based face recognition a thermal probe is matched against a gallery of visible faces corresponding to the real-world scenario, in a feature subspace. In cross-spectrum feature based approaches a function is learned to explicitly map the thermal features to the corresponding visible feature domain representation. The second category of approaches attempt to synthesize a visible-like face image from another modality such as NIR, thermal, or even polarimetric thermal input. These methods are beneficial because the synthesized image can be directly utilized by existing face recognition systems developed (i.e., trained) specifically for visible-based facial recognition. Therefore, using this approach one can leverage existing commercial-off-the-shelf (COTS) face recognition systems. In addition, the synthesized images can be used by human examiners for adjudication purposes.

2.1.2 Future Vision

The MWIR or LWIR imagery is ideal for night-time and low-light scenarios. However, the phenomenological differences between visible and thermal imagery, and the trade-off between wavelength and resolution (or pixel pitch) make matching visible and thermal facial signatures a daunting task. Current technology, based on cross-spectrum feature-based and cross-spectrum image synthesis methods have achieved significant performance in NIR-to-visible face recognition accuracy [2] and to some extent, for SWIR-to-visible face recognition accuracy [9]. However, MWIR-to-visible, or LWIR-to-visible face recognition systems still have a long way to go to achieve acceptable performance. Work is needed to narrow the modality gap between thermal and visible to improve the cross-spectral face recognition systems. One approach would be to integrate auxiliary information such as soft biometrics into the thermal-to-visible face recognition algorithms. For example, facial attributes (gender, color, ethnicity, etc.) can be used to alleviate the cross-modal gap by representing faces at a higher-level of abstraction. These facial attributes can in fact be automatically predicted [10] with high accuracy and be integrated with the Deep-Learning based face matchers [11]. Therefore, new seamless algorithms are needed to fuse soft biometrics with facial features for cross-spectral matching. The current cross-spectral databases are very small and overfitting the learning models with these small databases is a common issue. Large cross-spectral face databases are needed for experimentation with detailed facial attribute annotations. Another auxiliary information is the polarimetric measurement from the thermal face images. Polarimetric thermal imagery is represented by Stokes parameters: S0, S1, S2, where S0 is conventional thermal image, S1 captures the differences between the 0 degree and 90-degree polarization states, and S2 captures the difference between the 45 degree and 135-degree polarization states. It is possible to expand the capabilities of LWIR polarimetric imaging by generating three-dimensional (3D) images of human faces

reconstructed from single Stokes-vector images [12]. The reconstructed 3D-face can then be used to perform matching against a 3D-facial gallery.

Keywords

- Heterogeneous Face Recognition (HFR)
- Visible vs thermal face matching
- Facial attributes
- Cross-spectral face matching
- Thermal Polarimetric
- Soft biometrics

References

[1] S. Ouyang, T. Hospedales, Y.-Z. Song, X. Li, A survey on het-erogeneous face recognition: Sketch, infra-red, 3d and low-resolution, arXiv preprint arXiv:1409.5114.

[2] B. Klare, A. K. Jain, Heterogeneous face recognition: Matching NIR to visible light images, in: Pattern Recognition (ICPR), IEEE 20th International Conference on, 2010, pp. 1513–1516.

[3] F. Nicolo, N. A. Schmid, Long range cross-spectral face recognition: matching SWIR against visible light images, IEEE Trans-actions on Information Forensics and Security, Vol 7. No. 6, pp. 1717–1726, 2012.

[4] S. Hu, N. J. Short, B. S. Riggan, C. Gordon, K. P. Gurton, M. Thielke, P. Gurram, A. L. Chan, A polarimetric thermal database for face recognition research, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 119–126.

[5] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, V. M. Patel, Millimetre wave person recognition: Hand-crafted vs learned features, in: Identity, Security and Behavior Analysis (ISBA'17), 22-24 Feb. 2017.

[6] B. F. Klare, A. K. Jain, Heterogeneous face recognition using kernel prototype similarities, IEEE transactions on pattern analysis and machine intelligence, 35 (6) (2013) 1410–1422.

[7] J. Choi, S. Hu, S. S. Young, L. S. Davis, Thermal to visible face recognition, Tech. rep., Maryland Univ College Park (2012).

[8] S. Hu, J. Choi, A. L. Chan, W. R. Schwartz, Thermal-to-visible face recognition using partial least squares, JOSA A, 32 (3) (2015) 431–442.

[9] F. Nicolo, N. A. Schmid, Long range cross-spectral face recognition: matching SWIR against visible light images, IEEE Transactions on Information Forensics and Security 7 (6) (2012) 1717–1726.

[10] S. Ouyang, T. Hospedales, Y. zhe Song, and X. Li. 2014. Cross-modal face matching: beyond viewed sketches. Accepted by ACCV (2014).

[11] S. Mehdi Iranmanesh, B. Riggan, S. Hu, N. M. Nasrabadi, Coupled Generative Adversarial Network for Heterogeneous Face Recognition, Journal Image and vision computing, Volume 94, February 2020.

[12] A. J. Yuffa, K. P. Gurton, and G. Videen., Three-dimensional facial recognition using passive long-wavelength infrared polarimetric imaging. Applied Optics, Vol. 53, No. 36, 20 December 2014.

2.2 Technology Concept: Face Pose Invariant Features and Face Frontalization

2.2.1 Background

Unconstrained face recognition at extreme pose is an important and a challenging problem. The existing methods that address this pose problem can be generally categorized into two major categories. The first category aims to obtain pose-invariant embeddings [1, 2, 3]. The other aims to normalize (frontalization) face images [4, 5, 6, 7] to identity-preserved frontal views (as a preprocessing step), which can be directly used by any off-the-shelf face recognition systems without the knowledge of the recognition models. For the first category, deep metric learning [1] is a common way to achieve pose-invariant embeddings in a latent subspace. Due to the imbalanced distributions that characterize a long tail distribution of large pose faces, it is often difficult to achieve ideal pose-invariant embeddings across large pose variations. The second category is known as face frontalization, which resorts to computer graphics or deep learning to rotate profile faces to frontal views. Recently, deep learning-based methods have shown impressive capability on face frontalization [4, 5, 6, 7]. Early efforts [8] adopt the mean square loss to learn a deep regression model from paired training data. Recent methods have been proposed with novel network architectures [9, 10] or learning objectives [6, 11, 12]. Another interesting approach is to map the 2D face onto a canonical 3D model, this line of research date back to the 3D Morphable Model (3DMM) [13], which models both the shape and appearance as PCA spaces.

2.2.2 Future Vision

A challenge that persists is that of extreme poses – face-based models tend to breakdown on samples of faces that are viewed at extreme angles, pitches, and yaws. Furthermore, another major challenge is to design pose-correction algorithms or pose-invariant embeddings in unconstrained environments where there are more complex face variations, e.g., lighting, head pose, expression, self-occlusion in real-world scenarios.

Face normalization and embedding in the unconstraint environment. Current existing face normalization or embeddings techniques are developed under constrained environment and are based on supervised learning framework with the assumption of paired training data. More advanced techniques are needed to address pose correction and estimation for unconstrained environment with faces captured from surveillance cameras at distance. Developing algorithms that jointly address pose estimation, correction, and super-resolution is the next step in advancing face recognition systems. Extending these algorithms with to Multiview cameras [14, 15] or multiple pose-specific [3] methods are interesting line of research. Furthermore, combining concepts from both 3DMM and deep face recognition CNNs [5] to achieve high-quality and identity-preserving frontalization, we believe these approaches deserve further research.

Use of Multi-pose face normalization and embedding. Multi-view [14, 15] or multiple pose-specific [3, 16] methods are also used to obtain pose invariance features. The reason behind

using multi-pose face frontalization is that it can generate a better fronatlized face than the single view frontalization algorithms. For example, a multi-pose embedding technique can be used to aggregate all the off angle pose views into a robust feature embedding that can be used to reconstruct a single frontal face. The pros and cons of these multi-pose face fronatization approaches deserve further research.

Head pose estimation for extreme roll, pitch and yaw. Head pose is a 3D vector containing the angles of yaw, pitch, and roll. Estimating the head pose from an image essentially requires learning a mapping between 2D and 3D spaces. Methods utilizing more modalities such as 3D information in depth images or temporal information in video sequences are to be investigated. Algorithms still need to be developed that can estimate and correct for extreme yaw and pitch angles using faces at a very low-resolution (faces at a distance).

Keywords

- Face Normalization
- Face Frontalization
- Face Rotation
- Pose-invariant Feature Embeddings
- Face Landmark Detection
- Face Landmark Alignment
- Pose Estimation

RETURN TO CHART

References

[1] F. Taherkhani, V. Talreja, J. Dawson, M. Valenti, N. M. Nasrabadi, "PF-CpGAN: profile to frontal coupled GAN for face recognition in the wild," IEEE International Joint Conference on Biometrics (IJCB'20), Sept. 28- Oct. 1, 2020, Houston, Taxes.

[2] K. Cao, Y. Rong, C. Li, X. Tang, and C. L. Chen. Pose-robust face recognition via deep residual equivariant mapping. In IEEE Conference on Computer Vision and Pattern Recognition, 2018.

[3] I. Masi, S. Rawls, G. Medioni, and P. Natarajan. Pose-aware face recognition in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4838–4846, 2016.

[4] B. Y. R. H. Z. S. Yibo Hu, Xiang Wu. Pose-guided photorealistic face rotation. In CVPR, 2018.

[5] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker. Towards large-pose face frontalization in the wild. In ICCV, 2017.

[6] R. Huang, S. Zhang, T. Li, and R. He. Beyond face rotation: Global and local perception GAN for photorealistic and identity preserving frontal view synthesis. In ICCV, pages 2458–2467, 2017.

[7] J. Zhao, Y. Cheng, Y. Xu, L. Xiong, J. Li, F. Zhao, K. Jayashree, S. Pranata, S. Shen, and J. Xing. Towards pose invariant face recognition in the wild. In CVPR, 2018.

[8] J. Yim, H. Jung, B. Yoo, C. Choi, D. Park, and J. Kim. Rotating your face using multi-task deep neural network. In CVPR, 2015.

[9] Z. Zhu, P. Luo, X. Wang, and X. Tang. Deep learning identity-preserving face space. In ICCV, 2013.

[10] Z. Zhu, P. Luo, X. Wang, and X. Tang. Multi-view perceptron: a deep model for learning face identity and view representations. In NIPS, 2014.

[11] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker. Towards large-pose face frontalization in the wild. In ICCV, 2017.

[12] L. Tran, X. Yin, and X. Liu. Disentangled representation learning GAN for pose-invariant face recognition. In CVPR, 2017.

[13] V. Blanz and T. Vetter. A morphable model for the synthesize of 3D faces. SIGGRAPH, 1999.

[14] M. Kan, S. Shan, and X. Chen. Multi-view deep network for cross-view classification. In CVPR, 2016.

[15] X. Shao, X. Zhou, Z. Li and Y. Shi, "Multi-View Face Recognition Via Well-Advised Pose Normalization Network," in IEEE Access, vol. 8, pp. 66400-66410, 2020

[16] Z. Zhu, P. Luo, X. Wang, and X. Tang. Multi-view perceptron: a deep model for learning face identity and view representa-tions. In NIPS, 2014.

2.3 Technology Concept: Face Image Super-resolution

2.3.1 Background

Face super-resolution (FSR), a.k.a. face hallucination, refers to a class of computational methods for enhancing the spatial resolution of face images. FSR technology is deemed to a low-cost alternative to more expensive solution such as optical zoom. At the intersection of computational imaging and computer vision, FSR has found to be a valuable tool to facial biometric systems especially for extreme situations such as at a distance and from an altitude. Existing FSR technology can be classified into two categories: model-based and learning-based. Model-based FSR [1]-[3] include example-based [7], sparse coding based [8], and landmark-based [9]. For a review of face hallucination and generalized FSR, please refer to [1]-[3]. Learning-based FSR methods are represented by FSRNet [4] and its wavelet extension [5] as well as FSRGAN [4] and its unsupervised extension [6]. Earlier works on FSR have only considered the magnification ratio of ×2, ×3, and ×4. Recent trend in FSR has taken larger magnification ratio ×8 and beyond into account.

2.3.2 Future Vision

FSR in the wild. The majority of existing FSR is developed under the framework of supervised learning based on the assumption of paired training data. The simulated low-resolution (SLR) image is often obtained by artificially down-samplingdownsampling a high-resolution image. Such over-simplified assumption with degradation modeling does not match the characteristics of real low-resolution (RLR) images acquired in the real world. For example,

face detection on RLR images would have substantially less accurate bounding box location compared to downs-sampled high-resolution images. Such simulated-to-real gap has remained one of the long-standing open problems in face biometrics system including face recognition and FSR. How to extend FSR technology from supervised to unsupervised deserves systematic study. Some existing work (e.g., [6]) has proposed to use a generative adversarial network (GAN) to directly learn the degradation model associated with RLR images but with limited success. Our recent work [10] has leveraged the novel architecture of cycle-consistent GAN (CycleGAN) to FSR and demonstrated improved robustness when applied to RLR dataset. StyleGAN-based face image synthesis has also inspired the research into FSR techniques based on the extrapolation in the latent space of W+.

FSR and Face Recognition. Most FSR works aim to minimize some distance metrics defined with respect to the high-resolution images. However, the superresolved pixels might be just for a pleasure of eyes, but not more precise description of the identity information. Prior studies show that low-level vision tasks should be designed or trained end-to-end with high-level recognition tasks in order to benefit recognition [11]. With this mind, one may either bring the recognition loss into the FSR pipeline [13], or directly learn identity representations from RLR [12] where FSR could be an implicit module. The pros and cons of these approaches deserve future research.

Extreme FSR. Face recognition at long range (e.g., from hundreds to a thousand meters) or high altitude (e.g., from aerial platforms such as UAVs) has received increasingly more attention in recent years. IARPA has just initiated a new program named Biometric Recognition and Identification at Altitude and Range (BRIAR) to support this line of research. One of the primary technical challenges is the extreme low resolution (LR) of face imagery (face width of ~20 pixels) - it has been reported that the smallest spatial resolution for a human operator's discerning the facial identity information is around 18×24 pixels. Despite rapid advances in face recognition and super-resolution (SR), reliable extraction of face identity information from extreme LR face imagery such as Widerface (http://shuoyang1213.me/WIDERFACE/) has remained beyond the capability of current technologies. Recent advances in long-range and high-altitude biometric systems have renewed the interest in extreme face superresolution (FSR) - i.e., superresolving the resolution of face images by an extreme scaling factor (often larger than ×8). Extreme FSR in the wild has to address the challenges of preserving both face identity [13] and image quality while dealing with unknown degradation factors.

Keywords

- Face super-resolution (FSR)
- Simulated low-resolution (SLR)
- Real low-resolution (RLR)
- Generative adversarial network (GAN)
- Degradation modeling
- Face image prior

References

[1] Liu, C., Shum, H. Y., & Freeman, W. T. (2007). "Face hallucination: Theory and practice." International Journal of Computer Vision, 75(1), 115-134.

[2] Wang, N., Tao, D., Gao, X., Li, X., & Li, J. (2014). A comprehensive survey to face hallucination. International journal of computer vision, 106(1), 9-30.

[3] Jia, Kui, and Shaogang Gong. "Generalized face super-resolution." IEEE Transactions on Image Processing, 17, no. 6 (2008): 873-886.

[4] Chen, Yu, et al. "Fsrnet: End-to-end learning face super-resolution with facial priors." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

[5] Huang, H., He, R., Sun, Z., & Tan, T. (2017). Wavelet-srnet: A wavelet-based CNNcnn for multi-scale face super resolution. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1689-1697).

[6] Bulat, A., Yang, J., & Tzimiropoulos, G. (2018). To learn image super-resolution, use a GANgan to learn how to do image degradation first. In Proceedings of the European conference on computer vision (ECCV) (pp. 185-200).

[7] Park, Jeong-Seon, and Seong-Whan Lee. "An example-based face hallucination method for single-frame, low-resolution facial images." IEEE Transactions on Image Processing 17.10 (2008): 1806-1816.

[8] Yang, J., Tang, H., Ma, Y., & Huang, T. (2008, October). Face hallucination via sparse coding. In 2008 15th IEEE international conference on image processing, 2008, (pp. 1264-1267). IEEE.

[9] Yang CY, Liu S, Yang MH. Structured face hallucination. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2013 (pp. 1099-1106).

[10] Sidiya, Ahmed Cheikh, and Xin Li. "Style-based unsupervised learning for real-world face image super-resolution." In Recent Advances in Image Restoration with Applications to Real World Problems. IntechOpen, 2020.

[11] Scheirer, W., VidalMata, R., Banerjee, S., RichardWebster, B., Albright, M., Davalos, P., ... & Otto, C. (2020). Bridging the gap between computational photography and visual recognition. IEEE transactions on pattern analysis and machine intelligence.

[12] Xi Yin, Ying Tai, Yuge Huang, Xiaoming Liu , FAN: Feature Adaptation Network for Surveillance Face Recognition and Normalization, In Proceeding of Asian Conference on Computer Vision (ACCV 2020), Kyoto, Japan, Nov. 2020.

[13] H. Kazemi, F. Taherkhani, N. M. Nasrabadi, "Identity-Aware Deep Face Hallucination via Adversarial Face Verification," The Tenth IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS 2019), Sept. 23, 2019, Tampa, FL.

2.4 Technology Concept: Face Detection

tba...

2.5 Technology Concept: Gait Recognition

tba…

3.0 Biometric Modality Matching

Introduction

At the most abstract level, matching in pattern recognition is the process of measuring a degree of similarity between two patterns. The patterns of interest in this concept family are those that are specifically related to identity and/or the characterization of behavior for the purpose of detecting an action or inferring motive or intent. In the former, distinctive and measurable characteristics of individuals are used to identify or verify identity by comparing them to measurements stored previously for an individual. In the latter, distinctive and measurable characteristics of behavior are used to detect or classify a behavior by comparing the characteristics to those of a model for a targeted behavior of interest. Although technologies developed in this area can also be applied to animals, the primary focus of this roadmap will be applications related to human beings.

There are many general concepts related to pattern recognition and machine learning that form the backdrop for the technology used in biometric matching regardless of modality. Advancements in these areas have been critical to the current success and proliferation of biometrics in authentication and identification applications and will continue to have an impact on their development and adoption in society. For example, the success of matching is strongly correlated with the quality of the input data representation, which itself is affected by many factors. This implies a strong relationship between advancements in matching and advancements in all aspects of biometric sample characterization including, but not limited to: biometric sample quality, segmentation, feature extraction, classification and fusion. Many of these are touched on in other concept families; nevertheless, there are many overlapping research areas relevant to advancements in biometric modality matching.

3.1 Technology Concept: Physiological Biometrics

3.1.1 Background

Physiological biometrics refer to the use of physical measurements of the human body for the purpose of uniquely identifying or authenticating an individual. A reasonably comprehensive, but not exhaustive list is the following:

- Face
- Iris
- Hand Geometry
- Lips
- Periocular
- Finger-Knuckle Print
- Fingerprint
- Palm (Writer's Palm, Lower Palm)
- Ear Print
- Footprint
- Finger/Palm/Wrist Vein Mapping
- Retina

3.1.2 Future Vision

Major recent advancements have been made in face as a result of the confluence of three significant developments: advances in hardware computation, development of deep neural

network machine learning architectures and an explosion of available digital data. Large datasets for iris and fingerprints are also generally available, making it likely that deep learning approaches may result in advancement in iris and fingerprint matching technology as well, although further research is required to determine the upper limits of that potential [1, 2, 3]. Further advancements in all of these biometrics are expected as technology advances to harness the full potential of these developments. Numerous challenges still exist[4].

Procurement of large datasets is a significant issue for making advancements in these biometrics and others. More research is required into the potential of synthetic data generation to augment smaller datasets for training purposes [5, 6, 7]. Synthetic data generators have been found to be useful for fingerprint algorithm development and are being extended to face and iris [8, 9, 10]. Viable results have been obtained for segmentation, detection and landmark location, but this approach has yet to be proven effective for recognition training. Improved methods for data augmentation are also necessary to leverage the use of deep neural networks for biometrics with limited training data [11].

Benchmark datasets have been limited to face, iris and fingerprint, but their existence has aided in the advancement of commercial algorithms for those biometrics. Equivalent training and benchmark datasets would be valuable for other biometrics, specifically palm and multi-spectral face. Phone sensors using IR and NIR are coming online but progress will be hampered by limited training and benchmark data. Overall, methods to more efficiently use the little data that is available are necessary [12]. Potential variations on these are possible by studying either the passive detection or active reflection of various forms of radiation to obtain additional distinctive biometric representations based on physiological biometrics that could be useful for matching.

Most current applications of deep learning to biometrics concentrate on the training of feature extractors. As hardware continues to accelerate neural network processing, research into deep learning networks that include matching may provide additional boosts in matching accuracy.

There is a need for practical, conceptual frameworks that can accommodate the various requirements for matching engines. For example, how does one balance the need to be able to simultaneously discriminate between face images for unrelated individuals and family members or twins? How do differing performance criteria affect the design, deployment and integration of matching technologies in diverse operational environments? Can/should a single matching algorithm be expected to work optimally on biometric data from multiple image quality domains, or is there a need to uniquely address source biometric sample quality differences?

Cross-domain transfer learning [13, 14, 15, 16] is an important area for advancement, relevant across a range of domain differences, such as:

1. **Devices**
   Training and testing is often limited to a small number of devices. Transfer learning can be useful to generalize matching algorithms to devices not previously seen, but more research needs to be done. [17]

2. **Spectral modalities**
   Due to lack of training data, research into cross-spectral domain matching is of particular importance in cases where different types of radiation are used to generate or detect different biometric representations from the same physiological trait, e.g. infrared, polarimetric thermal, millimeter wave or depth representations for faces. There is a need to leverage transfer learning to enable robust cross-domain matching without requiring large domain specific training data.

3. **Biometric traits**
   Transfer learning has been used in a limited way to adapt object recognition neural networks to face recognition. The question is, can transfer learning be used to take advantage of these networks to make progress in others? The diverse nature of physiological biometrics may require the development of more specialized deep learning nets tailored to them to enable transfer learning to work more effectively. For example, new deep learning architectures may be required to deal optimally with certain biometric features i.e. the generic object detection networks may not be sufficiently targeted to the features needed to accurately match all biometrics. On the other hand, transfer learning may be more effective for networks designed for fingerprints but applied to iris (or vice versa), than networks originally designed for face.

Our understanding of the interaction between biometric sample quality and matching performance is limited primarily due to the ill-posed nature of the perceived problems [18, 19]. Matching involves a complex interaction between the biometric sample quality distribution of the gallery, the quality of the probe, the uniqueness of the gallery sample, the gallery size, and the inherent bias associated with the feature extraction and matching algorithms. A methodical and consistent approach to characterizing the effect of biometric sample quality on the matching process is still lacking, but is an active area of research [20, 21, 22].

Increased sensor resolution and decreased size/expense, as well as increased data transfer speeds and computing power will soon make practical the simultaneous acquisition of multiple biometrics (e.g. face/iris, periocular/iris, fingerprint/DNA, fingerprint/hand geometry, fingerprint/vein, retina/iris). As ubiquitous as the smart phone has become, there is no reason to think that the accessibility, diversity and use of sensors will not increase as well. Progress in smartphone acquisition of biometrics is accelerating but more research is required to ensure interoperability. This will likely require more sophisticated methods of fusion both at the score and/or feature level, along with the development of associated matchers [23, 24].

Multiple templates and template update can be powerful methods for improving the power of matching systems [25, 26]. More research is required into the tradeoffs between the operational deployment of frame selection algorithms, storage of multiple enrollment biometrics, fusion of enrollment biometrics and the incorporation of biometric sample quality into match strategies and decisions.

Finally, the ability to learn from a single instance is a powerful human skill, and one-shot learning algorithms try to mimic this special capability. Current deep learning algorithms require a significant number of examples to learn the multiple levels of representation necessary to recognize other individuals. Transfer learning can mitigate this to some degree. However, advances in one-shot learning that combine the two approaches could potentially yield powerful learning algorithms for many biometrics for which data is scarce [27].

Keywords

• deep learning • synthetic • data augmentation • transfer learning • big data • face
• fingerprint • iris • biometric sample quality • matching • one-shot

RETURN TO CHART

References

1. Ayanthi, J., Lydia, E.L., Krishnaraj, N. et al. "An effective deep learning features based integrated framework for iris detection and recognition". J Ambient Intell Human Comput 12, 3271–3281 (2021).

2. Sundararajan, K., Woodard, D.L. (2018). "Deep Learning For Biometrics: A Survey", ACM Computing Surveys, Volume 51, Issue 3, Article No.: 65, pp. 1-34.

3. Mehraj, H., Mir, A.H, "A Survey of Biometric Recognition Using Deep Learning", EAI Endorsed Transactions on Energy Web 03 2021 - 05 2021 | Volume 8 | Issue 33 | e6.

4. Shaheed, K., Mao, A., Qureshi, I. et al. "A Systematic Review on Physiological-Based Biometric Recognition Systems: Current and Future Trends", Arch Computat Methods Eng 28, 4917–4960 (2021).

5. Yanushkevich, Svetlana. (2006). "Synthetic Biometrics: A Survey", IEEE International Conference on Neural Networks - Conference Proceedings. 676 - 683.

6. Kar, A., Prakash, A., Liu, M.-Y., Cameracci, E., Yuan, J., Rusiniak, M., Acuna, D., Torralba, A. and Fidler, S. "MetaSim: Learning to Generate Synthetic Datasets", In ICCV, 2019.

7. Sankaranarayanan, S., Balaji, Y., Jain, A., Lim, S. N., and Chellappa, R., "Learning from Synthetic Data: Addressing Domain Shift for Semantic Segmentation", In CVPR, 2018.

8. Drozdowski, P., Rathgeb C., and Busch, C., "Sic-Gen: A Synthetic Iris-Code Generator," 2017 International Conference of the Biometrics Special Interest Group (BIOSIG), 2017, pp. 1-6.

9. Cappelli, R., Ferrara, M. and Maltoni, D. "Generating synthetic fingerprints", in Martin Drahanský, Hand-Based Biometrics: Methods and technology, IET, 2018.

10. Wood, E., Baltrušaitis,T., Hewitt, C., Dziadzio, S., Johnson, M., Estellers, V., Cashman, T.J., Shotton, J., "Fake It Till You Make It: Face analysis in the wild using synthetic data alone", arXiv:2109.15102, submitted to ICCV 2021.

11. Shorten, C., Khoshgoftaar, T.M. "A survey on Image Data Augmentation for Deep Learning", J Big Data 6, 60 (2019).

12. Adadi, A. "A survey on data-efficient algorithms in big data era". J Big Data 8, 24 (2021).

13. Kuzu, R. S., Maiorana, E., and Campisi, P., "Vein-based Biometric Verification using Transfer Learning," 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), 2020, pp. 403-409.

14. Aloweiwi, Aseel H., "Fingerprint Classification Using Transfer Learning Technique" (2021). Theses, Dissertations and Culminating Projects. 723.

15. Alaslni M.G., and Elrefaei, L.A., "Transfer Learning with Convolutional Neural Networks for Iris Recognition", International Journal of Artificial Intelligence & Applications (IJAIA) Vol.10, No.5, September 2019.

16. Ramos-Muguerza, E., Docio-Fernandez, L., and Alba-Castro, J.L., "From Hard to Soft Biometrics Through DNN Transfer Learning", Conference: IEEE 9th international conference on biometrics: theory, applications and systems, Los Angeles, California, USA, Oct. 2018.

17. Kandaswamy, C., Monteiro, J.C., Silva, L.M. et al. "Multi-source deep transfer learning for cross-sensor biometrics". Neural Comput & Applic 28, 2461–2475 (2017).

18. Grother, P. and Tabassi, E., "Performance of biometric quality measures," Pattern Analysis and Machine Intelligence, IEEE Transactions on 29(4), 531–543 (2007).

19. El-Abed, M., Charrier, C., and Rosenberger, C., "Quality assessment of image-based biometric information," EURASIP Journal on Image and Video Processing 2015(1), 1–15 (2015).

20. Hernandez-Ortega, J., Galbally, J., Fiérrez, J., Haraksim, R., and Beslay, L., "FaceQnet: Quality assessment for face recognition based on deep learning", IEEE International Conference on Biometrics ICB 2019, June 4–7, 2019, Jun. 2019.

21. Terhörst, P., Kolf, J. N., Damer, N., Kirchbuchner, F., and Kuijper, A., "Ser-fiq: Unsupervised estimation of face image quality based on stochastic embedding robustness.", In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 5651–5660, 2020.

22. Hua, F., Johnson, P., and Schuckers, S., "Utilizing automatic quality selection scheme for multi-modal biometric fusion," 2013 IEEE International Conference on Technologies for Homeland Security (HST), 2013, pp. 664-670.

23. Dinca, L. M., Hancke, G.P., "The Fall of One the Rise of Many: A Survey on Multi-Biometric Fusion Methods", Access IEEE, vol. 5, pp. 6247-6289, 2017.

24. Raju, A. S. and Udayashankara, V., "A survey on unimodal, multimodal biometrics, and its fusion techniques," International Journal of Engineering & Technology, vol. 7, no. 4, pp. 689–695, 2018.

25. X. Chen, M. Yu, F. Yue, B. Li and Y. Guo, "A Generic Template Update Framework: The Solution for Intra-Class Variations in Biometric Systems," in IEEE Access, vol. 7, pp. 145082-145094, 2019.

26. F. Yue and X. Chen, "Template Selection and Update for Biometric Recognition Systems With Nearest Neighbor Classifier," 2019 Chinese Control Conference (CCC), 2019, pp. 7797-7803

27. Sucholutsky, I., & Schonlau, M. (2021). `Less Than One'-Shot Learning: Learning N Classes From M < N Samples. Proceedings of the AAAI Conference on Artificial Intelligence, 35(11), 9739-9746.

3.2 Technology Concept: Soft Biometrics

3.2.1 Background

Soft biometrics are ancillary information such as facial measurements, color of the skin, height, gender, and ethnicity, which can be integrated to improve the overall performance of a primary biometric system. There are three categories of soft biometric modalities: global (gender, age, ethnicity), face (skin color, nose length, eye size, lip thickness), and body (arm length, chest width, height), which are considered more permanent modalities than attributes such as glasses and clothing [5]. Soft biometrics often lack the distinctiveness and permanence to sufficiently differentiate two individuals [8, 10]. Nonetheless, Soft biometrics as an alternative to traditional biometrics are emerging due to their independence, non-intrusiveness, semantics, and availability [5].

The application of soft biometrics through anthropometric and morphological measurements (head length, head breadth, length of middle finger, length of the left foot, and length of the cubit) dates back to Alphonse Bertillon's system in the 19th century that police could use to identify criminals via physical measurements, photography (mug-shot), and record-keeping [3]. The Bertillon system was later replaced by hard biometrics (fingerprint) due to the challenge in reliably identifying an individual via software biometrics. Soft biometrics can be used to identify individuals in the so-called hidden population (people who seek to remain hidden) or when biometric scanners are unavailable [6]. Soft biometrics can also be used to narrow down the search scope in the whole dataset [1]. However, due to lack of distinctiveness and permanence, soft biometrics were not trusted capable of being reliably used for user identification. Instead, they were integrated to enhance performance of a primary biometric system. For example, fingerprint as a primary biometric modality is combined with ancillary information of gender, ethnicity, and height resulting in a substantial improvement in performances [7, 8]. On the other hand, soft biometrics are recognized to be non-intrusive, independent, semantic, and easily available. Therefore, recent research

has emerged on developing soft biometric based standalone recognition systems. [9, 4] Soft biometrics can also help overcome disadvantages of a primary biometric, including acceptability (if users are not willing to provide biometric information such as face images), usability (difficult to interact), and circumvention (system resistant to spoofing). [1]

3.2.2 Future Vision

There are a few open challenges and recommendations for developing standalone soft trait-based recognition systems. Design and development of benchmark datasets – datasets available are used alone or in concatenation for user-recognition or evaluation of a soft biometric system. None of these datasets cover all the physical and behavioral modalities of humans. Mostly face and body are the primary focus. Quantitative annotations – annotations of the dataset are the next step after collection. The qualitative methods of annotation are good for short-term tracking or feature-based retrieval on a small dataset. But, for long-term tracking in an unconstrained environment, we need quantitative annotations. These include geometric or anthropometric measurements of soft traits. Feature selection – selecting features that are highly significant for recognition and retrieval can be challenging. Automated estimation of soft biometrics – any chosen soft trait to be used in recognition systems must pass through four factors, namely, feature correlation, permanence score, discrimination power, and attribute distance. Fusion – most of commonly used fusion techniques in standalone soft biometric-based recognition systems are feature-level (based on permanence and discrimination power) and modality level (based on the selected soft traits). Other challenges include recording environment, lighting conditions, random gaps between sessions, and lack of information about user demographics. [5]

Keywords

• Soft biometrics	• Hard biometrics	• Identification	• Permanence	• Distinctiveness
• Fusion	• Skin color	• Height	• Arm length	• Scar
• Gender	• Ethnicity	• Demographics	• Glasses	• Clothing

References

[1] Abdelwhab A., Viriri S. (2018), A Survey on Soft Biometrics for Human Identification, Chapter 3, IntechOpen. http://dx.doi.org/10.5772/intechopen.76021.

[2] Arigbabu OA, Ahmad SMS, Adnan WAW, Yussof S (2015), Recent advances in facial soft biometrics. Vis Comput 31(5):513–525.

[3] Bertillon system, (1883): https://en.wikipedia.org/wiki/Alphonse_Bertillon.

[4] Gurnani A, Shah K, Gajjar V, Mavani V, Khandhediya Y (2019), Saf-bage: Salient approach for facial soft-biometric classification-age, gender, and facial expression. In: 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp 839–847. IEEE.

[5] Hassan B., Izquierdo E., Piatrik T. (2021), Soft biometrics: a survey, Multimedia Tools and Applications, Springer. https://link.springer.com/article/10.1007/s11042-021-10622-8.

[6] Heckathorn D.D., Broadhead R.S., Sergeyev B. (2001), A methodology for reducing respondent duplication and impersonation in samples of hidden populations, J. Drug Issues 31 (2) (2001) 543–564.

[7] Jain AK, Dass SC, Nandakumar K. (2004), Soft biometric traits for personal recognition systems. In: Zhang D, Jain AK, editors. Biometric Authentication. Berlin/Heidelberg: Springer; 2004. pp. 731-738.

[8] Jain AK, Dass SC, Nandakumar K. (2004), Can soft biometric traits assist user recognition? in: Biometric Technology for Human Identification, 2004, pp. 561–573.

[9] Niinuma K, Park U, Jain AK (2010), Soft biometric traits for continuous user authentication. IEEE Trans Inf Forensic Secur 5(4):771–780

[10] Singh M., Singh R., Ross A. (2019), A comprehensive overview of biometric fusion, Information Fusion, Elsevier.

3.3 Technology Concept: Behavioral Biometrics

3.3.1 Background

Behavioral biometrics is the field of study that deals with the measure of uniquely identifiable patterns in human activities. These patterns of behavior are usually specific to the user and can be used for authentication, such as the rhythm with which a user types on his computer keyboard or mobile phone, the angle at which the phone is held when in use or the way a user walks in slow, normal and fast paces. In contrast, physiological biometrics involves the innate human characteristics such as face, fingerprints or iris patterns. Behavioral biometrics is a promising solution in financial institutions, businesses, government facilities and retail points for preventing account takeovers and fraud in a frictionless manner.

There are a plethora of behavioral biometric modalities but not all have been studied adequately. The metrics used for recording performances include, equal error rate (EER), area under the curve (AUC), half total error rate (HTER), false accept rate (FRR), False reject rate (FAR) and accuracy. Below are some modalities that have been relatively well studied.

1. **Keystrokes**

   Keystroke dynamics identifies users based on their typing rhythms and requires no extra hardware other than the keyboard readily available on a computer or smartphone. Furthermore, it is passive and non-intrusive. That is, it can run in the background in a frictionless manner without interfering with the user's activities. Keystroke dynamics can be classified into fixed-text and free-text. When users are constrained to type a predefined text, such as passwords, it is known as fixed-text, while free-text refers to

cases when users are allowed to type freely without restriction on what, when and how to type (e.g., writing an article on a topic of their interest). Sometimes, keystroke dynamics can be somewhere between fixed-text and free-text. This is known as semi-fixed text. The state-of-the-art for desktop fixed-text, semi-fixed-text and free-text are 9.6% [1], 0% [6] and 2.2% [7] EER respectively, while for mobile semi-fixed-text and free-text are 2.26% [6] and 9.2% [7] EER.

2. **Mouse Dynamics**
   Mouse dynamics is the process of identifying users based on their mouse operating behaviors. It is less intrusive and requires no special or additional hardware to capture mouse behavioral data. Mouse dynamics have two applications which are static mouse authentication and continuous mouse authentication. The former is when the mouse operations are used only once at a particular moment, as the case of user login. On the other hand, continuous mouse authentication continuously verifies the user's identity throughout the user's session. The state-of-the-art performance for continuous mouse authentication varies from 0.58% - 2.94% FAR and 0.12% - 2.28% FRR depending on the number of mouse actions used [8].

3. **Mobile Touch Gestures and Swipes**
   The touch gestures are prescribed shapes drawn on the mobile devices touch screen comprising of single or multiple strokes where each stroke is a series of successive numerical coordinates. Usable features of touch gestures include touch direction and duration, velocity and acceleration of movement. A study [2] produced 95.85% accuracy with 45 participants. Swipe is one of the dominant user actions on the touchscreen when people interact with their mobile devices. Like every other behavioral biometric modality, factors such as emotional state or injury can affect the accuracy of swipe. The state-of-the-art has performance ranging from 80% to 96% AUC [3].

4. **Walking Gait**
   This is based on the measurement and analysis of the way an individual walks or runs by using the acceleration signals produced by the gait recording device such as mobile phone. Smart phones have embedded sensors like accelerometer, gyroscope and magnetometer which can effectively measure gait characteristics with no additional cost. Factors such as device orientation change, uneven ground, possible injuries, fatigue or footwear can affect the accuracy of walking gait when used for authentication [4]. Performance for gait varies based on where the sensor is placed (waist, left-side, front-side, pocket etc.), the type of sensor used (camera, smartwatch, smartphone, floor sensor, accelerometer etc.) and the number of users in the study. The state-of-the-art performance ranges from 0.17% to 2.27% EER for wearable sensors and 1.23 to 4.07% EER for smartphone.

5. Speaker Recognition

Speaker recognition is the identification of a person from characteristics of voices. It is used to answer the question "Who is speaking?". It is one of the most convenient and accessible behavioral biometric modalities due to the abundance of devices equipped with a microphone, such as smartphones. Applications of speaker recognition include forensics, surveillance, user authentication etc. Identification/ authentication performance may be affected by the acoustic mismatch induced by varied environments and devices of the same speaker. The state-of-the-art performance is 0.9% EER for normal-normal speech, and 17.8% - 27.3% EER for normal-whispered speech [5]. Note that speaker recognition is different from speech recognition, which is the ability to recognize spoken words or commands from speech.

3.3.2 Future Vision

Behavioral biometrics lacks benchmark datasets to compare algorithms and generalize performance claims that can be trusted. Many studies are impossible to replicate as researchers have reported performances based on their private datasets. In addition to the lack of benchmark datasets, the number of users/participants involved in each study is relatively small to generalize.

Keywords

- Behavioral biometrics
- Implicit authentication.
- Continuous authentication
- Usability
- Keystroke dynamics
- Mouse dynamics
- Speaker recognition
- Mobile touch gestures
- Mobile swipes

References

[1] Killourhy, Kevin S., and Roy A. Maxion. "Comparing anomaly-detection algorithms for keystroke dynamics." In 2009 IEEE/IFIP International Conference on Dependable Systems & Networks, pp. 125-134. IEEE, 2009.

[2] Yang, Yafang, Bin Guo, Zhu Wang, Mingyang Li, Zhiwen Yu, and Xingshe Zhou. "BehaveSense: Continuous authentication for security-sensitive mobile apps using behavioral biometrics." Ad Hoc Networks 84 (2019): 9-18.

[3] Wang, Xiao, Tong Yu, Ole Mengshoel, and Patrick Tague. "Towards continuous and passive authentication across mobile devices: an empirical study." In Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks, pp. 35-45. 2017.

[4] Stylios, Ioannis, Spyros Kokolakis, Olga Thanou, and Sotirios Chatzis. "Behavioral biometrics & continuous user authentication on mobile devices: A survey." Information Fusion 66 (2021): 76-99.

[5] Vestman, Ville, Dhananjaya Gowda, Md Sahidullah, Paavo Alku, and Tomi Kinnunen. "Speaker recognition from whispered speech: A tutorial survey and an application of time-varying linear prediction." Speech Communication 99 (2018): 62-79.

[6] Wahab, Ahmed Anu, Daqing Hou, Stephanie Schuckers, and Abbie Barbir. "Utilizing Keystroke Dynamics as Additional Security Measure to Protect Account Recovery Mechanism." In ICISSP, pp. 33-42. 2021.

[7] Acien, Alejandro, Aythami Morales, John V. Monaco, Ruben Vera-Rodriguez, and Julian Fierrez. "TypeNet: Deep Learning Keystroke Biometrics." arXiv preprint arXiv:2101.05570(2021).

[8] Hu, Teng, Weina Niu, Xiaosong Zhang, Xiaolei Liu, Jiazhong Lu, and Yuan Liu. "An insider threat detection approach based on mouse dynamics and deep learning." Security and Communication Networks 2019 (2019).

3.4 Technology Concept: Behavior Identification

3.4.1 Background

Modality matching in the context of behavior identification in this document will be limited to classification of human behavior that depends on an analysis of individual behavior for the purpose of determining intent. The following is a non-exhaustive list of some examples:

1. Micro-expressions
2. Emotion detection
3. Body expression recognition
4. Human activity detection

Although somewhat tangential to biometrics, these patterns are distinctly human and can be used to characterize subgroups of individuals by behavior, analogous to how soft biometrics can be used to characterize subgroups of individuals based on physical characteristics. Like soft biometrics, they may have components to them that may be individually distinctive, but not independently sufficient for identification.

For example, video systems can view individuals before attacks and collect information on individuals who frequent potential attack sites, providing a baseline for identifying individuals engaged in pre-execution activities. Fig. 1 below shows the relationships between various components of a threat monitoring system that combines information from multiple levels of analysis of human behavior, with various degrees of individuation [1].

3.4.2 Future Vision

Depending on the application, a potentially serious problem arising from the use of this technology is the extremely high false detection rate. For example, in threat detection scenarios, there is a need for imminent behavior threat detection, but high false detection rates can not only severely inconvenience the innocent but can also divert and distract scarce and valuable resources that could be better used elsewhere. Improving detection but at the same time determining optimal methods to combine such information without incurring such errors is critical to this progress.

Although normal facial expression recognition is now considered a well-established field with accuracies exceeding 90%, automatic recognition of micro-expressions in contrast, is still relatively new with many challenges. One of the challenges is detecting the micro-expression of a person accurately from a video sequence. Micro-expressions are typically very subtle and of short duration, making detection of micro-expressions a difficult task. Detection is even more difficult when the video clip consists of spontaneous facial expressions and unrelated facial movements, e.g. eye-blinking, opening and closing of mouth, etc. Additional challenges include inadequate features for recognizing micro-expressions due to their low differential intensity. Although deep learning methods would be expected to be particularly amenable to solving this problem, lack of data hampers significant advancement. Annotated datasets exist only for several hundreds of subjects [2]. In contrast to face recognition, the dependence on facial features makes 3D synthetic face generation combined with computer animation a viable alternative for supplying training data, but more research is required.

From another perspective, as data on human behaviors are increasingly digitized, the volume of potentially analyzable data increases accordingly, and manual annotation and analysis becomes impractical or impossible. Unsupervised methods for training may be necessary for future advancement of this technology [3, 4].

Emotion recognition algorithms that fuse multiple parameters seem to perform much better than ones that infer emotional state simply from facial expressions alone. Of course, checkpoints or other security environments are dynamic locations where it is difficult to capture high-resolution video, but improvements in surveillance technologies will inevitably yield that type of information, making such data fusion more feasible. One of the most frustrating deficiencies in the labeling of data associated with emotion and body expression research is the lack of

consistency in the taxonomy used for naming various states. This results in considerable fragmentation of the data and makes transfer learning techniques difficult [5].

It has been noted by many researchers that there is considerable complementarity in different modalities. Unfortunately, research in multimodal emotion recognition remains rather scarce and simplistic. The little research that exists mostly focuses on simple fusion techniques using shallow representations of the body and face or body and speech. Even though all methods report important improvements over unimodal equivalents, this potential remains largely unexplored [5].

Observation of indicators representative of mental states holds promise for the detection of deception and other behaviors. For example, polygraph testing, and other measures of peripheral nervous system response such as heart rate, heart beat, blood pressure, electroencephalograms (EEGs), vocal stress, and facial expression and micro-expression analysis have the potential for detecting deception and/or hostile intent. Two primary problems with using physiological indicators are non-specificity (the indicators may stem from many causes some of which may be benign) and individual differences (the observables that indicate attack or deception differ markedly across individuals, which may require matching against individual-centered baselines)[6, 7].

New technologies using non-contact electric field sensors also allow some physiological features to be observed without physical contact, sometimes at some distance, and sometimes covertly or surreptitiously, as with using heat-sensitive cameras to detect capillary dilation and blood flow to the face and head. There is some evidence of improving detection of deception or imminent action by an individual if baseline information is available for that specific individual ahead of time.

Keywords

• micro-expressions • body gestures • body expression recognition • emotion
• human activity detection

References

1. Davis, Paul K., Walter L. Perry, Ryan Andrew Brown, Douglas Yeung, Parisa Roshan, and Phoenix Voorhies, "Using Behavioral Indicators to Help Detect Potential Violent Acts: A Review of the Science Base", Santa Monica, CA: RAND Corporation, 2013.

2. Oh, Y., See, J., Ngo, A.C., Phan, R.C., Baskaran, V.M., "A Survey of Automatic Facial Micro-Expression Analysis: Databases, Methods, and Challenges", Front Psychol. 2018; 9: 1128. Published online 2018 Jul 10.

3. Fabian Benitez-Quiroz, C., Srinivasan, R., & Martinez, A.M. (2016). "Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild", In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5562–5570.

4. Heilbron, F. C., Escorcia, V., Ghanem, B. and Niebles, J. C., "ActivityNet: A large-scale video benchmark for human activity understanding," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 961-970.

5. Noroozi, F., Corneanu, C. A., Kamińska, D., Sapiński, T. Escalera, S. and Anbarjafari, G., "Survey on Emotional Body Gesture Recognition," in Journal of IEEE Transactions on Affective Computing, vol. 12, no. 2, pp. 505-523, 1 April-June 2021.

6. Lee, D., Zhan, P., Thomas, A., and Schoenberger, R.B., "Shape-based human detection for threat assessment", Proc. SPIE 5438, Visual Information Processing XIII, (15 July 2004).

7. Qing Lei, Ji-Xiang Du, Hong-Bo Zhang, Shuang Ye and Duan-Sheng Chen, "A survey of vision-based human action evaluation methods", Sensors, vol. 19, no. 19, pp. 4129, 2019.

8. Nguyen, D. T., Li, W., Ogunbona, P. O., "Human detection from images and videos: a survey", Pattern Recognition 51 (2016) 148– 175.

Concept Family 4.0: Data Analytics and Fusion

4.1 Technology Concept: Real-Time High Speed Intelligent Graph Data Analytics

4.1.1 Background

Applications of graph theory and social networking are useful for understanding the texture and context of interactions in the social networks of individuals who are potential adversaries, and possibly for predicting their behavior. Col Glen Voeltz (1), in his monograph on "The Rise of IWar" outlines in detail the growing flood of both classified and open source information available to the intelligence community that is amenable to social network (read: graph) analysis and identification of potential adversaries. Advancements in these areas will most likely be associated with "Big Data" technologies, because the volume, the variety, and the velocity of the data requires large, distributed databases and streaming data technologies. Voeltz points out the importance of biometrics in these advancements. It has become clear that the commercial advanced analytics in the spheres of both machine learning and deep learning, combined with the new graph analytics that have come to characterize customer behavior analytics are uniquely adaptable and parallel to the new requirements in Identity Warfare (IWar) and intelligence fusion analysis, and track well with the new biometrics capture and correlation capabilities to include biometric biographical data as well (2).

4.1.2 Future Vision

Within an intelligence-based knowledge graph system as described in Figure 1, multi-modal

encounter-based biometric data stores serve as data sources providing information on encounter locations and times as well as specific communications linked to identities. Unconstrained face and voice biometric data collection provide significant value.

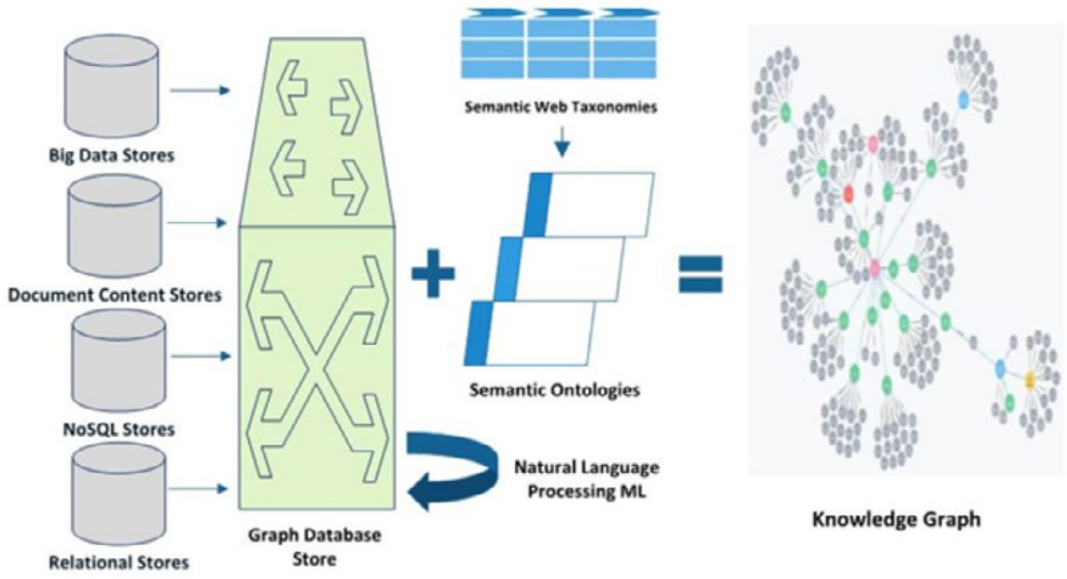

Figure 1. Components of Knowledge Graph Systems

Current Knowledge Graph solutions, generically, can use a variety of structural components because they are multi-modal in terms of the origin and the structure of the data they use for their construction. The most common components are numerous data stores, including but not limited to document stores, relational stores, NoSQL stores, and other big data/Hadoop stores which can hold a variety of data types. These stores generally have a translation layer to a graph store (itself considered a NoSQL database). This graph store can then be informed by a semantic/ontological layer which defines semantic hierarchies, and which itself is defined by taxonomies built from web systems. These layers can often be augmented or combined with Lucene/Solr/Elastisearch layers which index and provide faceted search capabilities. The combination of these ontologies linked to the structures of the underlying graph data makes the construction of derived knowledge graphs possible, as depicted in Figure 1.

The translation layer between the different stores each require different and extensive preprocessing, and in the more advanced Knowledge Graph systems, there is extensive Natural Language Processing (NLP) using machine learning- for example, deep learning for classification using Convolution or Recursive Neural Networks (CNNs and RNNs) and machine learning techniques such as LDA (Latent Dirichlet Allocation) for topic analysis. This NLP could also involve the construction of Word2Vec, GloVe, LDA2vec, and other even more advanced pre-analyzed semantic spaces using a language corpus specific to the subject area (3,4).

In order to implement solutions to solve very large graph problems many coinciding problems arise. It is clear that for most large knowledge graph systems, the main problem comes before the graph data. The system needs to be:

- Near real-time (refreshed often and from various stores)
- High throughput (to transfer large amounts of data between systems)
- Store large quantities of information (billions of rows yielding billions of nodes and edges)
- Data will be in various structured and unstructured formats (text, images, key/pair data)

This means an operational Big Data system with a full streaming and publish and subscribe layer is required. There are several different tiers of big data systems, each requiring a different architecture. Operational systems are built for speed and volume of data throughput and speed of the data stores and on-board data processing systems to stream data in from external sources and between system components.

Based on the current state of the art, three areas for future work are identified.

1. Current technologies require network fusion to be done iteratively for different data sources, with each being added through a different transformation and then iteratively assimilated into a graph database. Transformations are needed which allow data to be associated across data sources through data pipe integration, before it enters the graph database, with an objective to speed graph integration.
2. Current technologies link graphs with knowledge hierarchies derived and set in RDF and OWL frameworks, which are difficult to generate and maintain, and very difficult to modify with rapidly and continuously streaming data. Work is needed to replace these knowledge hierarchies with knowledge models generated for NLP hierarchical topic modelling, with an objective to build a more tightly integrated and quickly adaptable framework for providing appropriate knowledge frameworks to knowledge graphs. The object will be to find adaptable analytics frameworks that simplify and accelerate integration and building of semantic knowledge bases, using cutting edge NLP technologies.
3. Current technologies leave large gaps where solving the last mile of the problem, they assume that all the hard work of integration and data preparation and delivery is already accomplished. Work is needed to evaluate the most advanced and appropriate Big Data streaming models with intermediate model processing and integration of machine learning have to offer to solve the hard issues leading up to that last mile. This work will examine the competing Big Data frameworks and newest options, like in-stream processing and analytics with automated semantic hierarchy production and update.

Keywords

- Graph data analytics
- Voice recognition
- Unconstrained face biometrics collection
- Encounter based biometrics
- Natural Language Processing
- Unconstrained voice biometrics collection

References

1. Colonel Glen Voeltz, "The Rise of IWar: Identity, Information, and the Individualization of Modern Warfare - Terrorism in Iraq, Afghanistan, Use of DNA, Biometrics, Forensics, Palantir, and Facial Recognition", US Government Dept of Defense, (January 2017).
2. Taylor Anderson, "Biometrics ATP 2-22.85: Multi-Service Tactics, Techniques, and Procedures for Tactical Employment of Biometrics in Support of Operations (MCRP 3-33.1J, NTTP 3-07.16, AFTTP 3-2.85, CGTTP 3-93.6)", Dept of Defense (August 2016).
3. Yoav Goldberg. Neural Network Methods for Natural Language Processing. Morgan and Claypool Publishers, Synthesis Lectures on Human Language Technologies. (2017).
4. Jason Brownlee, "Deep Learning for Natural Language Processing". Machine Learning Mastery Series (2017).

5.0 Concept Family: Security, Privacy, Perception

5.1 Technology Concept: Security

5.1.1 Background

Cognitive security concerns protecting our cognitive process from deliberated attacks. Media forgeries, including DeepFakes, can be considered as a way to hack our biological perceptual system and decision making process. Thus they pose risks to cognitive security with a rippling effect to the overall security of biometric systems. In addition, synthesized impersonating media (images of faces, irises, finger prints, voices, and videos) can be exploited to gain access to biometric systems in the form of a rebroadcast attack. There are overlappings between multimedia forensics and likeness detection.

The current efforts in Multimedia Forensics heavily tilt towards detection, which formulate the problem as a binary classification between real and fake media. The state-of-the-art detection methods are based on DNN classifiers trained on large datasets. Although these methods have demonstrated promising performance on various benchmark datasets, detection alone is not adequate to combat DeepFakes [1].

- Detection methods only operate post mortem after DeepFakes emerge and cause damages.
- Anti-forensic attacks can deliberately mislead the detection method to make classification errors by hiding traces of DeepFakes, especially those based on deep neural networks.
- Detection methods usually do not provide direct evidence of DeepFakes beyond the Yes/No answer.
- Little information is revealed about the generation process of DeepFakes using detection methods.

- The current detection methods are inefficient in running time, and cannot be used for real-time detection for applications such as liveness detection.

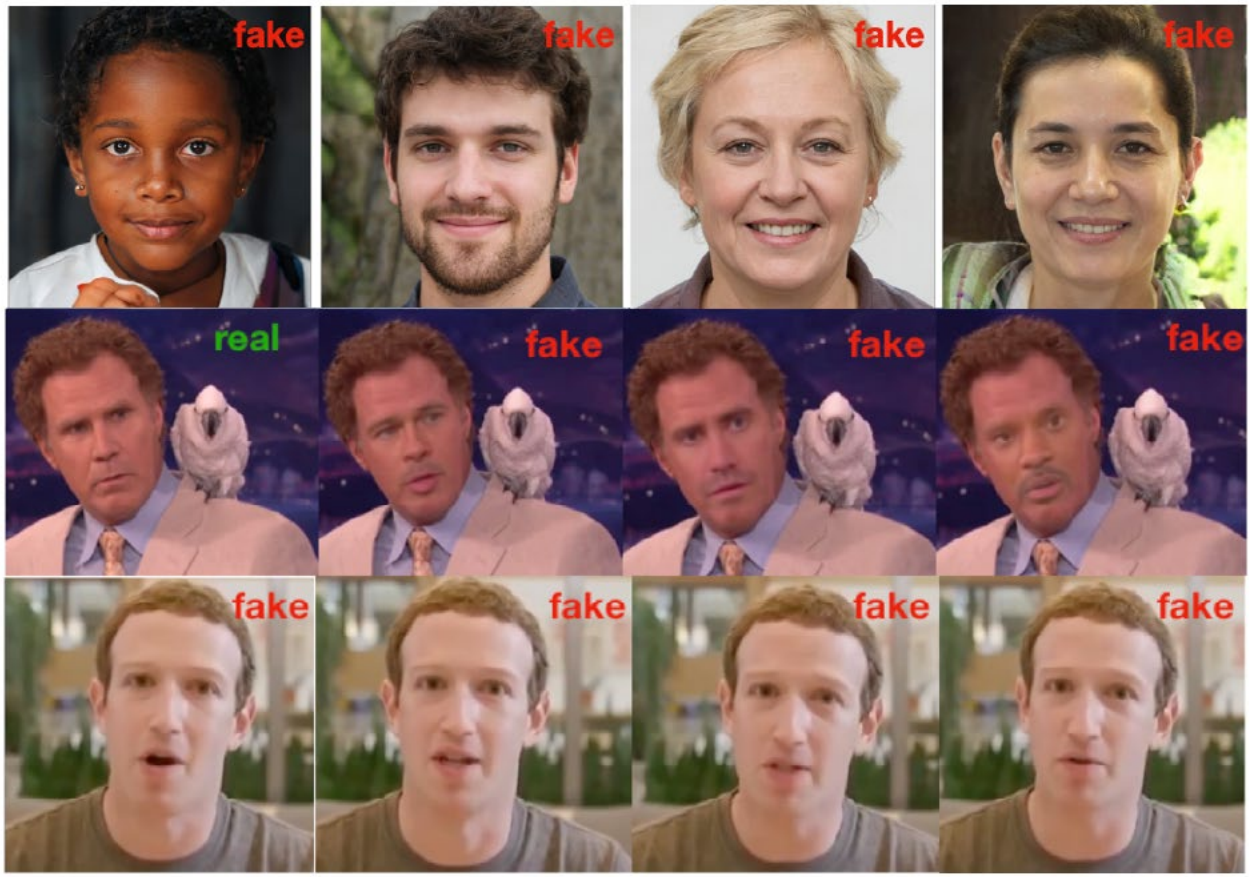

Figure 1. Various types of media forgeries that pose threats to the cognitive security and biometric systems. (Top): GAN synthesized faces. (Middle) Face-swaps. (Bottom) Lip-syncing.

5.1.2 Future Vision

There is no doubt that we are going to see more media forgeries and DeepFakes in the coming years, with better visual qualities, cheaper and easier to create, and taking new forms. Therefore the detection and other counter-technology need to catch up with the pace. Given the limitations of passive detection methods, there are needs to develop more active approaches. Instead of waiting for the DeepFakes generated and then identifying them using detectors, One such direction is a more active approach by adding impercep- tible traces to the would-be training data. Such traces should be learned by the DeepFake generation models trained using the tainted data, and can be extracted later from the DeepFakes created using the tainted model. The traces provide definite evidence of synthesis [2,3]. The other direction is to develop a proactive approach to obstruct DeepFake generation by sabotaging the training process. This is achieved by poisoning the would-be training data to disrupt

the critical pre-processing steps including face detection and landmark extraction. The poisoned data will lead to reduced efficiency and low quality (occasionally total failure) of the synthesized DeepFakes. The poisoned data can attack the models directly [4] or important pre-processing steps in model training [5].

CITeR is uniquely situated to combine multimedia forensic research with biometric research, and the cross-pollination between the two areas is expected to lead to new effective methods.

Keywords

• Cognitive security • data poisoning • backdoor attack

RETURN TO CHART

References

[1] Siwei Lyu. DeepFake Detection: Current Challenges and Next Steps. In International Workshop on Media-Rich Fake News (MedFake) in conjunction with ICME, London, UK, 2020.

[2] Run Wang, Felix Juefei Xu,Qing Guo,Yihao Huang, Lei Ma, Yang Liu, and Lina Wang. Faketagger: Robust safeguards against deepfake dissemination via provenance tracking. In ACM Multimedia, 2021.

[3] Ning Yu, Vladislav Skripniuk, Sahar Abdelnabi, and Mario Fritz. Artificial gan fingerprints: Rooting deepfake attribution in training data, ArXiv 2020.

[4] Chin-Yuan Yeh, Hsi-Wen Chen, Shang-Lun Tsai ,and Sheng-De Wang. Disrupting image-translation-based deepfake algorithms with adversarial attacks. In Proceedings of the IEEE/CVF Winter Confer- ence on Applications of Computer Vision (WACV) Workshops, March 2020.

[5] Pu Sun, Yuezun Li, Honggang Qi, and Siwei Lyu. Landmark breaker: Obstructing deepfake by disturbing landmark extraction. In IEEE Workshop on Information Forensics and Security (WIFS), New York, NY, United States, 2020.

5.2 Technology Concept : Privacy

5.2.1 Background

The privacy concerns over many biometrics stem from the fact that biometric templates in their raw format, such as fingerprints, iris, and face, are permanent, unique identifiers that can be stolen and exploited by adversaries to launch spoof or replay attacks, or used to cross-match the individuals among multiple databases without informed consent (function creep), thus violating data privacy protection laws. [1] The current practice of using cryptographic encryption/decryption to secure the raw biometric templates is limited

in that stolen cryptographic keys may potentially lead to large-scale, catastrophic data and privacy breaches. One way to address this limitation is to store the encrypted template and decryption key in a secure environment that the owner directly possesses, such as a smart card or a secure chip, but its applicability is limited to verification (1:1) rather than identification (1:n). In addition, two broad categories of template protection algorithms (feature transformation vs biometric cryptosystems) have been proposed in the literature to achieve the requirements of noninvertibility, revocability, and non-linkability for the generated templates but without compromising recognition performance. However, significant research advances are needed before practical acceptance of such algorithms, including design of invariant biometric representations with high entropy and guaranteed performance, independent benchmarking, and practical solutions ensuring revocability and nonlinkability of protected templates. [2]

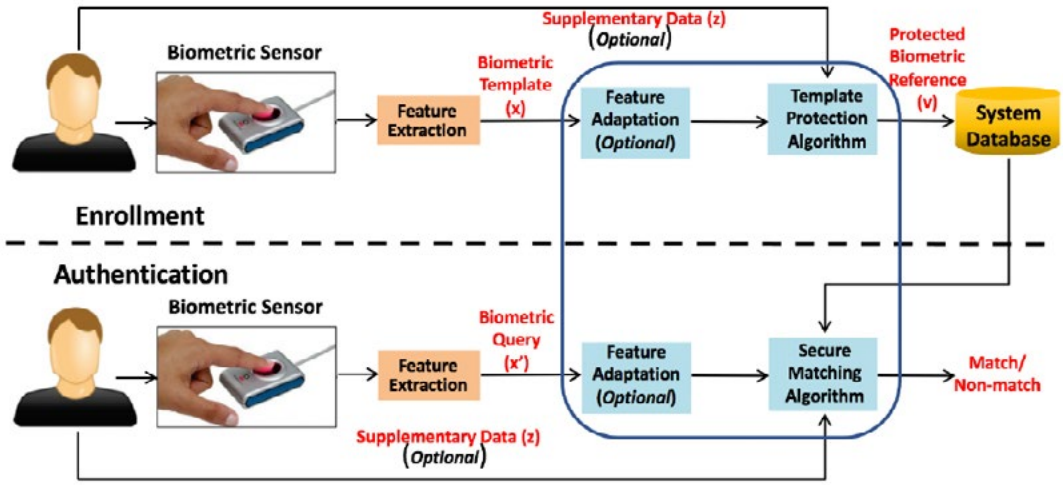

Figure 1. Architecture of biometric systems integrated with template protection.

5.2.2 Future Vision

Self sovereign identity [3] can become a major design paradigm that offers strong privacy protection, especially in verification-based scenarios, such as access control and account-based online financial transactions. In this regard, the FIDO standard can be promising to enjoy widespread industry adoption.

Secure biometric data sharing will continue to be an important design concern for identification-based sceneriors, such as border control, covert surveillance, and forensics, as it would make sense for agencies to share access to their databases in order to be more effective in accomplishing their mission. As noted above, significant research is needed to mature template protection algorithms (feature transformation vs biometric cryptosystems)

as well as to evaluate their relative merits and design tradeoff against the existing encryption-based solutions. CITeR can play a major role in connecting affiliates with researchers to jointly initiate such studies.

Keywords

• Biometric template • template protection • feature transformation • biometric cryptosystems

RETURN TO CHART

References

[1] P. Campisi, Ed., Security and Privacy in Biometrics. Springer, 2013.

[2] K. Nandakumar and A. K. Jain, "Biometric Template Protection: Bridging the performance gap between theory and practice," in IEEE Signal Processing Magazine, vol. 32, no. 5, pp. 88-100, Sept. 2015, doi: 10.1109/MSP.2015.2427849.

[3] Reza Soltani, Uyen Trang Nguyen, Aijun An, "A Survey of Self-Sovereign Identity Ecosystem", Security and Communication Networks, vol. 2021, Article ID 8873429, 26 pages, 2021. https://doi.org/10.1155/2021/8873429.

5.3 Technology Concept: Perception

5.3.1 Background

Although people are now far more likely to encounter and willingly engage with biometric recognition technologies in the course of their day than just a few years ago, automated biometric recognition—what it is and how it is used—remains a technological mystery to most people. Against this backdrop, perceptions of the technology are all-too-easily (negatively) influenced by news reports of wrongful arrests predicated on erroneous facial recognition matches, exposés about elaborate biometrically-enabled surveillance systems, and revelations of questionable biometric data collection practices. The security and efficiency gains expected to be realized through the adoption of automated biometric recognition are thus at risk of being lost in some cases for want of a better understanding of the technology and its many and varied applications.

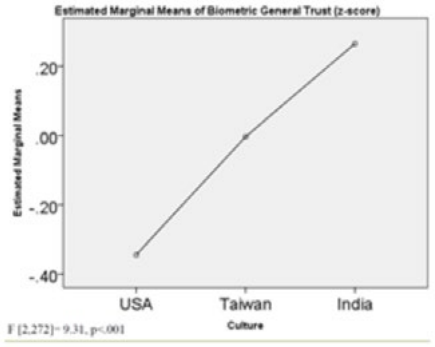

Table 5. Trust dimensions with unique dimensions and shared dimensions

	Interpersonal	General Technology	Automation	Information Technology
Dimensions	Behavior, Trustworthiness, Ability, Integrity, Benevolence, Reliability, Faith	Ability, Ease of Use, Reliability, Compatibility, Usefulness, Risk, Behavior, Functionality, Helpfulness, Attitude, Intension, Faith	Attitude, Usefulness, Ease of Use, Subjective norm, Reliability, Faith, Behavioral Intention, Workload, Complexity	Faith, Ease of Use, Reputation, Risk, Attitude, Usefulness, Subjective norm, Behavior
Unique Dimensions	Integrity, Trustworthiness, Benevolence	Helpfulness, Reliability	Workload, Complexity	Reputation
Shared Dimension	**Behavior, Ability**	**Ease of use, Attitude, Ability, Behavior**	**Attitude, Ease of use**	**Ability, Behavior, Ease of use**

Figure 1. Example results from a survey showing differences in trust related to biometric recognition technology in different parts of the work (left) and dimensions of consideration (right) [1]

5.3.2 Future Vision

In our future vision, we expect that when a person uses a biometric system, it is clear to the user what the system is doing and how it might be different from other applications. It is also clear what the biometric data is being used for and the privacy principles that apply. Decision/policy makers understand differences in biometric technologies and applications in order to develop reasonable laws, policies, and practices. When a biometric decision is made, mechanisms are in place to explain the automated decision, e.g. explainable AI.

CITeR's role is to educate decision makers and the general public on biometric technology, applications, architectures, and databases. Explicit development of laws and policies is outside the scope of CITeR, but rather we envision that we provide educational material to ensure policies are based on accurate understanding of the technology.

Education content includes aspects such as what databases are queried in a biometric match (e.g. criminal databases, drivers license photos, passport photos) and what images are used to query a database (e.g. video from CCD after a crime is committed, people in a public square). Also this understanding will be useful for developing safety mechanisms to control things like when a database is queried, when is a database search allowed, e.g. which crimes allow a search, whether a court order is needed, and whether a system is auditable Additionally, CITeR research includes studies of perception and trust in biometric recognition including focus groups, surveys, etc.

CITeR also has a role in terms of developing methods to explain a biometric decision, i.e. explainable AI. Biometric recognition has largely moved to Deep Neural Networks (DNNs). However, due to their complexity and multiple layers of abstractions, it is not easy to obtain a clear, interpretable relationship between the inputs and outputs of a DNN. The field of eXplainable AI aims to address this issue by explaining and representing this relationship in a human understandable terms [2]. This explainable relationship plays a crucial role in

various use cases such as decision making involving both humans and DNNs, verifying and debugging generalization of the model, efficient retraining of models and improving transparency to prevent unexpected behavior and unintended discrimination.

Keywords

- Perception
- Biometric education
- Privacy
- Explainable AI
- Surveys
- Focus groups
- Trust in technology

References

[1] Semnani-Azad, Z., Chien, S.Y.J., Forster, Y., Schuckers, S. and Gan, H., 2019, January.Development of Trust Measure in Biometric Technology. In Proceedings of the Annual Hawaii International Conference on System Sciences.

[2] Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, pp.82-115.

[3] Mia and Sophia Explain Biometrics, CITeR Biometric Recognition Educational Video Series, https://www.youtube.com/channel/UCwSnCVSl4AJQ37iDZ3v3SIw

---

6.0 Concept Family: Fairness, Demographic Differentials, Distinctiveness

6.1 Technology Concept: Fairness

Introduction

In Machine Learning (ML) and consequently in Biometrics as well that is relying on ML, fairness issues arise from the analysis of figures of merit (e.g. accuracy) in specific demographics groups (e.g., gender, ethnicity, race, revenue levels, or any covariate in general) and the observation that operational conditions originally estimated cannot be reproduced in those. The large-scale deployment of such systems in so many different scenarios raises the debate about its fairness and its impact on our lives.

6.1.1 Background

Many criteria to assess and address fairness in ML problems have been proposed over the years, each phrasing the problem differently. Recent work hypothesizes that most of the criteria described in the literature boil down into three major categories of conditional independence, and they are independence, separation, and sufficiency.

Independence (also called demographic parity) requires that a classifier must be independent of the sensitive attributes, i.e., the prediction must be identical no matter which of the sensitive attributes is present. This is also addressed as demographic parity or statistical parity. Separation (also addressed as equalized odds) explicitly acknowledges that the classification output might be correlated with the sensitive attributes which might be a desirable property in some applications. Sufficiency requires that the same decision threshold is used and independent sensitive attributes such as demographic factors.

In the biometrics literature, aspects of fairness are recently addressed for some biometric traits. For instance, the Face Recognition Vendor Test (FRVT) has a special report addressing demographic effects in face recognition (FR), mostly observing the effect of race and gender on more than 100 Commercial-off-the-Shelf (COTS) systems. Recent work observed that the "Other Race Effect", which is well-known in humans, can also be observed in FR algorithms – FR systems developed in Asia are more accurate with Asians than with Caucasians, and vice-versa. Furthermore, this work also highlights the role that image quality plays in fairness in FR systems. It was observed that four COTS systems presented higher verification rates for Caucasians than East Asiatics, using the Ugly set from The Good, The Bad, and The Ugly dataset. The FRVT report also raised similar observations. Studying race, consistently higher False Match Rates (FMR) with African American cohorts compared with Caucasians were observed using two COTS systems. The impact of age in FR was a topic of an extensive study. Recently, Wang et al. introduced the Racial Faces in the Wild dataset, a subset of the MSCeleb-1M whose identities are organized in four different races: Caucasians, Black, Indian and Chinese. Using such data to test the independence criterion, the authors regularized different deep neural networks at training time by minimizing the Mutual Information between the FR classifier and the demographic attributes. In other biometric traits, biases toward gender were also observed in the periocular region of the face. For instance, previous work demonstrated that several periocular recognition systems perform better with male subjects than with females. The NIST SRE is the most relevant benchmark for speaker recognition, and along with past editions, it consistently evaluates error rates looking at gender cohorts.

6.1.2 Future Vision

In recent work, Pereira and Marcel proposed a novel figure of merit, the Fairness Discrepancy Rate (FDR)[1], to assess fairness in biometric verification systems and address how the biometrics community considers fairness in operational conditions. The vast majority of works in biometric verification that assess demographic discrepancies does so by comparing ROC curves, DET Curves and/or the area under those curves. For instance, the area under the ROC curve is used to assess how fair a biometric verification system is under different demographic groups. ROC curves measure the True and False Positive Rates (TPR and FPR, respectively) trade-offs. Although this seems sensible to assess demographic discrepancies, it assumes that the verification decision threshold is demographic-specific. Assessing fairness with those standard figures of merit gives the impression that biometric

verification systems are fair while they are not. The FDR assesses fairness by verifying that both FMR and FNMR are equally separable between different demographic groups under the same decision threshold and allows to compare how fair two biometric verification systems are with respect to different demographic attributes for single decision thresholds.

Keywords

• Demographic differentials • Demographic parity • Bias Assessment and Mitigation
• Responsible datasets (synthetic datasets) • Self-evaluation and certification scheme of fairness

6.2 Technology Concept: Demographic Differentials in Operational Systems

Introduction

Algorithm-based demographic differentials are often amplified in operational systems as a result of other factors, including the quality of the biometric sample capture equipment, the quality of the biometric reference, overall system design (e.g., camera placement relative to the subject), and environmental conditions. The demographic effects that have been documented through controlled testing, such as that conducted by NIST, thus can be even greater in real-world applications.

6.2.1 Background

A biometric recognition system's performance is a function of many factors. As awareness of the problem of algorithmic bias has grown, so too has attention on these other critical elements of the biometric recognition process. Research carried out by the United States

Department of Homeland Security, for example, has shown that the quality of the image acquisition system in facial recognition systems "…can strongly affect (magnify or eliminate) measured differences in algorithm accuracy across demographic categories." Elsewhere, an analysis of an automated border control system revealed a significant variation in matching performance between different kiosks located in the same airport, despite the fact that all of the kiosks were identical and using the same algorithm. On inspection, it was discovered that variations in natural light were influencing the performance of the individual kiosks. The importance of non-algorithmic factors is known to extend to other biometric modalities as well. A study of a fingerprint recognition system yielded evidence that the observed differences in demographic performance might be a consequence of the quality of the equipment used to capture the biometric sample.

6.2.2 Future Vision

[1] http://publications.idiap.ch/index.php/publications/show/4463

While the problem of algorithmic bias has been extensively documented in relation to facial recognition technology, considerable work remains to be done to understand the full scope of the problem in relation to other biometric modalities. That means there is still a great deal to be learned about how algorithmic biases impact different types of biometric recognition systems and how any demographic effects can be mitigated or eliminated through improvements or changes in other parts of the system. In the near-term, therefore, additional research effort should be directed towards understanding all of the factors, along with their interactive dynamics, that affect overall system performance. Assuming the problem of algorithmic bias remains statistically relevant, the longer-term objective would be to develop systems that can compensate for any algorithm-based variances.

Keywords

- Algorithmic Bias
- Demographic Differentials/Effects
- Biometric System Design
- Operational Performance

6.3 Technology Concept: Distinctiveness

Introduction

The "distinctiveness" of a biometric trait refers to how unique that trait is to an individual. In some cases, the term "individuality" has been used to suggest that a biometric trait can be unambiguously associated with a single individual. However, it is generally agreed that determining the individuality of a biometric trait may not be easy, but that computing the distinctiveness or uniqueness of a given biometric trait might be more practical and relevant [refs].

6.3.1 Background

Consider a fingerprint image, F, from which a set of n minutiae points are extracted. What is the probability that k of these n minutiae points match with k minutiae points from another fingerprint image corresponding to a different finger? More generally, how many other fingerprints have the same configuration of k minutiae points as fingerprint F does? This is sometimes referred to as the probability of random correspondence. This can be viewed as a measure of the distinctiveness of a fingerprint [refs].

A similar question has been posed for the face and iris modalities with respect to the features commonly used to represent them [refs]. Thus, the distinctiveness of a biometric trait depends upon (a) the specific features used to denote it, and (b) the "model" used to describe the features.

To be added: describe "features" and "models", problems due to identical twins and how identical twins impose an upper bound on recognition accuracy], Overuse of the word "unique" - cautionary.

6.3.2 Future Vision

The distinctiveness of a biometric trait has practical utility in forensic applications where the nature and quality of the evidence has to be evaluated. It is also beneficial in developing more effective biometric recognition algorithms that account for individual-specific features. Further, knowing the distinctiveness of an individual's trait can help in better addressing the problem of dictionary attacks (such as MasterPrint attack).

Some of the research problems in this area include:

1. Developing and validating methods for estimating the distinctiveness of a biometric trait with respect to some features
2. Determining how this information can be incorporated into the biometric comparator for improving recognition accuracy
3. Large-scale data collection representing different demographic, cultural, and social groups
   - Determining if degree of uniqueness varies across these groups
4. Studying the impact of data quality on the distinctiveness of a trait
   - Factors impacting quality must be considered
   - Does "degree of uniqueness" vary across age
5. Designing fusion techniques based on degree of uniqueness of a person's trait

Keywords

• Uniqueness • Evidential Value • Individuality Models • Dictionary Attacks
• Identical Twins

RETURN TO CHART

7.0 Concept Family: Other, extension to areas beyond biometrics

7.1 Technology Concept: Optics and Biometrics

Just as the prints on a person's fingers and palms form a unique biometric, so too do the vein structures underneath the skin—each individual has a unique blood vessel pattern in their fingers and palms that others cannot mimic. Traditionally, optical methods scanning the venous structures in a person's hand use NIR light (750 nm to 1500 nm) and polarization. NIR light can penetrate through biological tissue of about 3 mm where it's absorbed at a specific rate by the deoxygenated blood in veins, creating a darker shadow in the image that effectively maps the venous vasculature in the palm or the finger.

There are two main methods of palm-vein imaging: reflection or transmission. The reflection method, which relies on light absorption and reflection to reconstruct the vascular image, is most common. In

this method, the illuminating component (the light source) and the capturing component (the sensor or camera) are on the same side of the target—the front. The transmission method, however, places the target in the middle, sandwiched between the illuminating and capturing components. Of course, this requires a stronger light penetration and is therefore less common.

These biometric systems frequently use CCD cameras to capture the palmar or finger veins. NIR CCD cameras provide higher resolution; however, the equipment can be expensive. Most vein-imaging modalities use LED light sources emitting in the 800-nm-to-900-nm range for an optimal image. Adding polarization or NIR filters boosts the resolution of the vascular structures in the biological tissue. Polarization transforms a light wave to only vibrate in a single plane, like how polarized sunglasses block glare. In the same vein, polarizer lenses placed in between the light source and the camera filter out light reflections and scattering from the surface of the skin to provide better contrast between the vascular structure and the surrounding tissue.

Moving along the biometric three-step process, the captured venous image is then uploaded into an image-processing algorithm for feature extraction followed by authentication. The biometric authentication protocol includes two modes: the registration and enrollment mode, which registers the individual's biometric into the database, and the authentication process, which verifies the individual's identity. Once enrolled, the authentication system can then recognize the individual with the biometric authentication protocol as a legitimate or an adversary drone.

Just as 3D facial recognition is more secure than 2D systems, 3D vein imaging biometrics is harder to spoof than 2D techniques. Photoacoustic tomography is a burgeoning imaging modality that can overcome the light-scattering limits in human tissue that make 3D biometrics difficult to obtain. In this technique, a laser illuminates and is absorbed by blood vessels in human tissue, creating an ultrasonic shockwave, which is subsequently detected by an ultrasonic transducer array. Then, using a physics-based acoustic source localization algorithm, a 3D image of the blood vessels is reconstructed. Just last year, our team at NEC Laboratories America and the University of Buffalo, USA, demonstrated accuracies greater than 99% and false acceptance rates as low as 0% with this system.

Key advantages of optical-based biometric technologies are non-invasiveness, harmlessness and the ability to obtain biometrics at a distance. Research that will see rapid growth in the near future will address the persistence of those advantages amidst requirements of the "new normal," such as contact-freedom, preservation of privacy and equity and anti-spoofing. That includes the biometric application of advanced imaging modalities, such as photoacoustic tomography, optical coherence tomography and lidar.

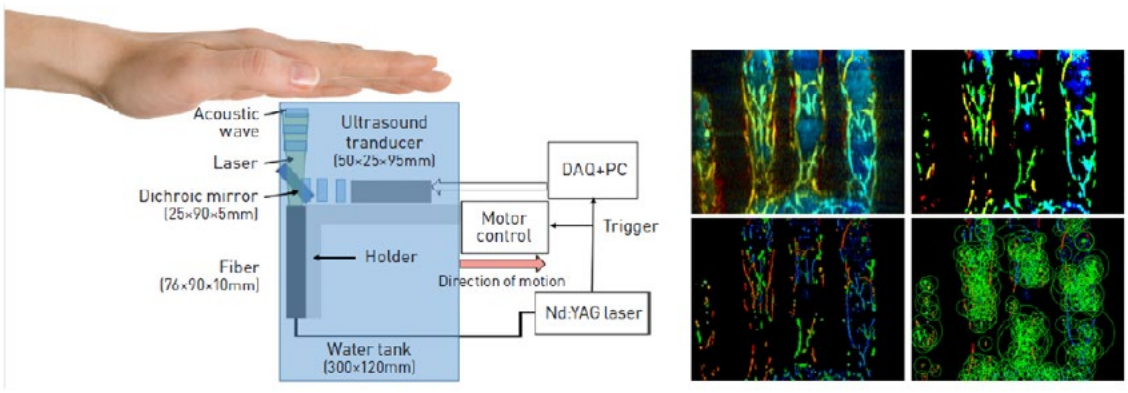

Figure 1. Example of vein imaging biometrics system and resulting images.